# General Disclaimer

## One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

# GLOBAL CROP PRODUCTION
# FORECASTING DATA SYSTEM
# ANALYSIS

7.9 - 1 0. 1 9 8

CR - 161199

CONTRACT   NAS8-32408
JANUARY
1978

## FINAL REPORT

ECOSYSTEMS INTERNATIONAL INC

P.O.Box 225
Gambrills, Maryland 21054

# GLOBAL CROP PRODUCTION FORECASTING

# DATA SYSTEM ANALYSIS

PETER A. CASTRUCCIO, HARRY L. LOATS, DONALD G. LLOYD

FINAL REPORT

JANUARY 1978

CONTRACT NAS8-32408

PREPARED FOR

THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

MARSHALL SPACE FLIGHT CENTER

MARSHALL SPACE FLIGHT CENTER, ALABAMA 35812

BY

ECOSYSTEMS INTERNATIONAL, INC.

P.O. BOX 225

GAMBRILLS, MD 21054

# TABLE OF CONTENTS

-i-

# TABLE OF CONTENTS (cont'd)

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF FIGURES   (cont'd)

# FOREWORD

The objective of Contract NAS8-32408 was to assess the characteristics required by the Data Systems of future OSTA Earth-Observation systems of the next decade.

It specified investigations in the six application areas Agriculture, Water Resources, Land Use, Global Weather, Severe Storms and Marine Weather. Reported here are the results of the analysis of Agricultural Data Systems. The authors wish to express particular thanks to Mr. Harvey Golden and Dr. George McDonough of Marshall Space Flight Center for their enlightened leadership: and to Dr. Charles T. Paludan, Dr. Sanford W. Downs, Jr. and Dr. Robert Jethro of Marshall for their many valuable suggestions.

Significant portions of this work were developed in prior efforts under the leadership of Dr. Louis Walter and Mr. William Stroud of Goddard Space Flight Center, and Mr. Leonard Jaffe of NASA Headquarters. The author's thanks extend to them also.

Dr. James C. Harlan provided unstinting cooperation in the difficult task of analysis of the spectral data bank gathered by his organization in Bushland, Texas.

## 1.0  SUMMARY, RESULTS AND CONCLUSIONS

The theme for the Data System task in Agriculture, chosen in concert with the Marshall Space Flight Center, was "Global Crop Production Forecasting." Selection of a specific theme within the broad area of Agriculture was motivated by the desire to lend focus to the task; simultaneously, Marshall's directives were to provide results sufficiently general as to be applicable to other agricultural uses of remotely sensed data.

The analysis herein reported applies to a generic agricultural observation system:  LANDSAT D was used as a benchmark against which to compare quantitative results.

Two principal scenarios motivate Global Crop Forecasting:  "Promotion of the US Balance of Trade" and "Promotion of Global Freedom from Want." The former was chosen as the more realistic motivator of US efforts.

The crops to be surveyed were the major US export staples.  They were selected from the cumulative ten-year export history of the US. The rationale for selecting countries candidate for survey excluded non-producers of specific target crops:  it included overproducers, i.e. competitors to the US for exports, and marginal producers which steadily or periodically import US agriproducts.

The selection methodology yielded the staple crops wheat, soybeans, corn, cotton, tobacco, rice, grain sorghum, and 24 countries growing these commodities.  Of these crops, tobacco and cotton were excluded

from further consideration because they do not constitute alimentary staples.

Substitution of crops was found to be negligible in human food, important in animal fodder. This added to the list of commodities the small grain fodder crops barley, oats and rye.

Crop production is the product of the two essentially independent factors of acreage and yield. Since measurement of the latter from remotely sensed data is as yet in the experimental stage, the study concentrated on the measurement of acreage.

The user requirements were assessed as being approximately equal to the accuracy performance of current conventional crop acreage inventory systems. The corresponding systems specifications for accuracy range from 98% for US wheat to 92%, one sigma, for foreign forage crops.

The performance of OSTA data systems in acreage mensuration was assessed from analysis of the performance achieved by ERTS and LANDSAT experimenters. Average inventory accuracy was found to be of order 80% in the time frame 1972-74, improving to approximately 86% between 1974 and 1978. These levels of performance were found to fall short of the user requirements by a factor of at least 2, at most 5.

This finding led to investigate the data system from end-to-end, in order to seek the reasons for the shortcoming, and to determine which data system drivers required improvement.

The analysis began at the data systems front-end, i.e. at the interface between the natural phenomena determining crop spectral reflectance and the sensor. To this effect, several of the fine-grained

calibrated spectra from the LACIE supersites were analyzed in detail. It was concluded that unique crop spectral signatures do not appear to exist; that the values of the reflectances fluctuate significantly, even within the same field; that the information contained in the differential amplitudes of spectral bands is in and by itself insufficient to achieve discrimination accuracies satisfying the user requirements; and that the driver of discrimination is the relative average amplitude, or contrast ratio, between the mean spectral amplitudes, brough about by the marked differences of mean reflectance which accompany the natural stages of phenologic development.

These findings led to the development of a theory of radiometric discrimination employing the mathematical framework of the theory of discrimination between scintillating radar targets. The theory indicates that the functions which drive accuracy of discrimination are the contrast ratio between targets, and the number of samples, or pixels, observed.

The theoretical results led to three primary consequences as regards the data system: 1) agricultural targets must be imaged at correctly chosen times, when the relative evoluation of the crop's development is such as to maximize their contrast; 2) under these favorable conditions, the number of observed pixels can be significantly reduced with respect to wall-to-wall measurements; 3) the remotely sensed radiometric data must be suitably mixed with other auxiliary data, derived from external sources.

Analysis of the optimal timing requirements in the presence of naturally occurring cloud cover led to the conclusion that the current

configuration of earth-observing platforms operating in the optical and IR spectral ranges need modification as to orbits and swath widths, and could probably substantially benefit from the addition of active microwave sensors. Analysis of the optimal sampling requirements indicated that the number of pixels imaged and processed drives the data system's operating costs: the use of a sampling scheme in heir of the current wall-to-wall imaging mode could reduce the system's recurring costs by up to a factor of two. Analysis of the role of the auxiliary data indicates that a most important contribution to the system's price-performance would be a world-wide, uniform, photographic-type coverage of relatively good resolution. Such coverage need not be provided in real time: its refresh rate could range upwards of three years. These findings can form the basis for considerable enhancement of the price/performance of future OSTA agricultural observation systems. In particular, the RBV could become the source for the auxiliary photocoverage. Additional savings are possible by on-board processing of the data, consisting of the addition of short-term storage.

An additional systems savings can result from relaxation of the multispectral scanner's sensitivity. This is brought about by the presence of fluctuations, found to exist around the reflectances of agricultural targets. Within the overall system's error budget, of which the "scene fluctuation" is the primary driver, sensor NE $\Delta\rho$'s of order 64 levels appear adequate.

The resulting 4:1 reduction of sensor sensitivity can be traded off against lower-cost sensor implementation: or, it can be exploited to increase the sensor's geometric resolution. It can be shown

theoretically that the tradeoff in favor of performance would improve
LANDSAT D's resolution to approximately 15 meters.

The theoretical results were simulated on the Marshall Data System
Dynamic Simulator (DSDS). The results confirmed the need for modifying
the deployment of earth-observing platforms, and the savings accruing
from observing a reduced number of sample picels. The simulation neither
confirmed nor disproved the theory, because the DSDS had not reached its
full Stage III configuration, thus could not, at the time, accept inputs
from live scenes. The theory was however tested independently by per-
forming regional crop inventories in selected regions of the world, with
good agreement between results and predictions.

## 2.0 SCENARIO

### 2.1 Significance of Global Crop Survey and Forecast

The activities of estimating and forecasting agricultural production are of major importance virtually everywhere in the world.

Inaccurate forecasts cause agricultural producers to make erroneous production decisions: they distort optimal inventory carryovers. The net result is lowered social efficiency through wasted efforts. Improved forecast accuracy reduces this waste motion, hence produces a net increase in social welfare.

Precise forecasts obtained sufficiently early permit the formulation of national plans and policies best suited to cope with surplus or scarcity. Good predictions allow agricultural producers to take remedial actions aimed at increasing output, reducing costs, or taking shelter against declining prices caused by overproduction. The earlier this information is available, the greater the possible spectrum of ameliorative actions and the larger the potential benefits.

The worth of agricultural survey and forecast activities is evidenced by the efforts and resources which are being devoted to it worldwide. Table 2-1 presents a conservative estimate of the expenditures by world governments for agricultural data gathering. Additional moneys are expended yearly by commodity dealers, traders, Chambers of Commerce, international organizations, other private and public enterprises [1].

TABLE 2-1

YEARLY WORLDWIDE EXPENDITURES FOR CROP SURVEY/FORECAST (1977)

| REGION | SURVEY/FORECAST COST MILLION U.S. $, 1977 |
|---|---|
| U.S. (DOMESTIC & FOREIGN) | 40 |
| CANADA | 2.4 |
| WESTERN EUROPE | 48.2 |
| EASTERN EUROPE | 8.9 |
| LATIN AMERICA | 2.5 |
| AFRICA/MID-EAST | 2.1 |
| ASIA EXCL. USSR, PRC | 4.5 |
| OCEANIA | 1.6 |
| WORLD | 110.1 |

The world's important crops number in the hundreds. Most important in terms of market value, quantity and caloric content are the staple crops.

Figure 2-1 shows that grain crops are the most significant: they provide over half of mankind's food. [2], [3].

## 2.2 World Scenarios

The configuration and characteristics of a Data System for global agricultural forecast are a function of the surface area, geographic location, number of crops to be surveyed.

Two principal scenarios determine the selection of the countries and crops candidate for survey:

Figure 2-1

# PANORAMA OF THE WORLD'S PRODUCTION
# OF STAPLE CROPS



Million Tons

GRAINS
U.S. SHARE

**BY WEIGHT**
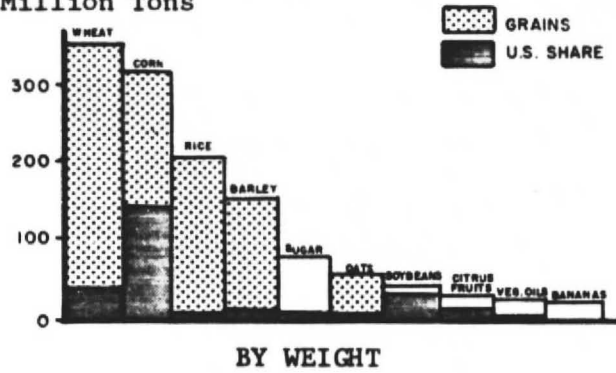


45 BILLION U.S DOLLARS (1978)

**BY VALUE (AT INTERNATIONAL EXPORT PRICES)**



WCY's

I WCY - AVERAGE YEARLY CALORIC REQUIREMENT
OF WORLD POPULATION (4 BILLION @ 2,500
CAL/DAY) =
3,650 TERA CALORIES

**BY CALORIC CONTENT**
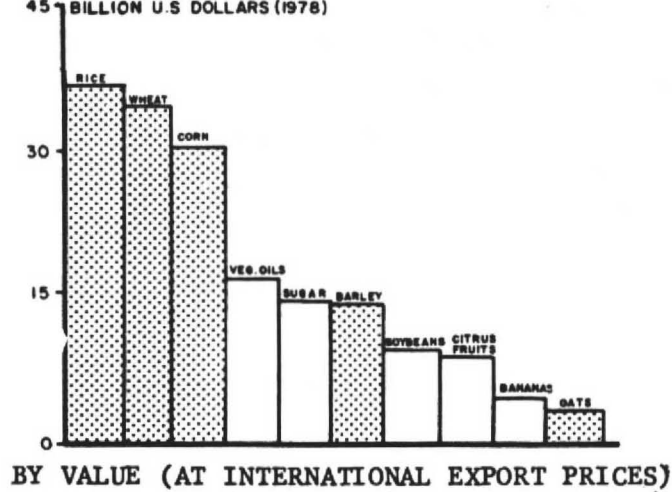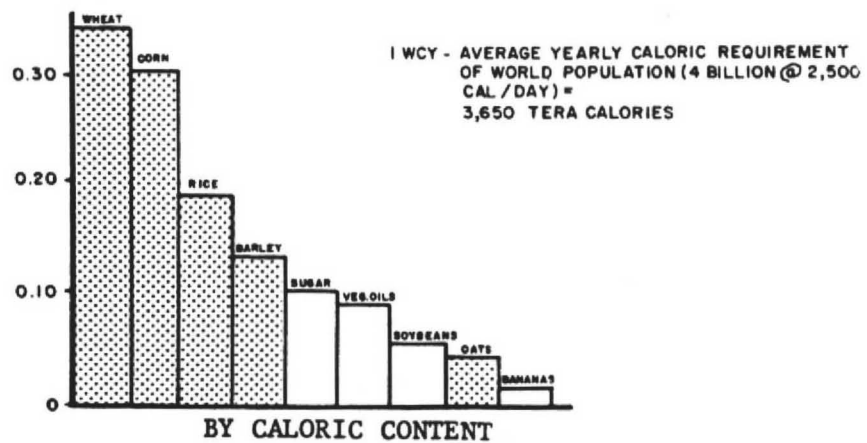
1. As seen from the U.S. national outlook

2. As seen from the outlook of the international community, i.e.
   the United Nations.

The U.S. national goals have essentially remained unchanged throughout U.S. history. They are:

- Maximize the citizen's well-being

- Foster the U.S. balance of trade

- Optimize U.S. foreign policy

The United Nations goal is essentially the humanitarian objective:

- Foster world peace

- Promote global freedom from want

## 2.3 The U.S. Scenario

In the agricultural arena, the three U.S. goals are interrelated.

Promotion of the balance of trade implies essentially "sell more, at more advantageous prices."

Citizen's well-being, or consumer satisfaction, means: "but without unduly increasing domestic prices for U.S.-produced commodities, and maintaining favorable prices for imported agricommodities."

The opportunity for a significant foreign policy role arises if the U.S. can trade food in exchange for political or socieconomic considerations.

This study focused on the first goal, i.e. maintenance and promotion of a favorable balance of trade.

## 2.4  Derivation of the Crops to be Surveyed

Agricultural products have always played a major role in the U.S. export picture.  In the mid-nineteenth century they represented approximately 80% of total U.S. exports.  Although the agricultural share of total U.S. exports has since declined to 25%, the gross dollar value of agricultural exports has increased from $150 million in the 1870's to over $25 billion, in current dollars.

U.S. agricultural exports fluctuate yearly, in response to domestic and foreign alternations of production shortfalls and surpluses.  To smooth out these flucutations, the export values should be integrated over a sufficiently long period of time.  Table 2-2 ranks the U.S. agriexports based upon their 10-year cumulative value [3].

TABLE 2-2

### U.S. GROSS AGRICULTURAL EXPORTS, 10-YEAR CUMULATIVE VALUE
### 1964-1974, CURRENT $ MILLION, FOB

| COMMODITY | ABSOLUTE VALUE | | RELATIVE VALUE | |
|---|---|---|---|---|
| WHEAT | $18,527 | | 18% | |
| SOYBEANS (FOR FEED) | 14,750 | | 15% | |
| CORN | 13,873 | | 14% | |
| ANIMAL & ANIMAL PRODUCTS OF WHICH | 10,892 | | 11% | |
| FATS OILS & GREASE | | 2,803 | | (3) |
| MEAT & MEAT PRODUCTS | | 1,866 | | (2) |
| OIL SEEDS OF WHICH | 9,090 | | 9% | |
| SOY OILCAKE & MEAL | | 4,299 | | (4) |
| COTTONSEED OIL | | 660 | | (1) |
| COTTON | 6,525 | | 6% | |
| TOBACCO | 6,186 | | 6% | |
| FEED GRAINS, EXCLUDING CORN OF WHICH | 4,021 | | 4% | |
| GRAIN SORGHUM | | 3,004 | | (3) |
| RICE | 3,684 | | 4% | |
| SUBTOTAL | 87,548 | | 86% | |
| OTHER | 14,481 | | 14% | |
| TOTAL 10-YEAR EXPORTS | $102,029 | | 100% | |

Table 2-3 synthesizes the information of Table 2-2. It excludes the animal products, for the following reasons:

- The consistency of the animal inventory is not directly amenable to survey from Satellite Remote Sensing.

- The diet of animals is composed of grains, pasture and forage in varying proportions in differing countries. The feed grains are already included in the listing of Table 2-3.

### TABLE 2-3

#### GROSS U.S. EXPORTS OF STAPLE CROPS, 10-YEAR CUMULATIVE VALUE
#### 1964-1974, CURRENT $ MILLION, FOB

|  | VALUE | PERCENT |
|---|---|---|
| SOY AND SOY PRODUCTS | 19,049 | 21 |
| WHEAT | 18,527 | 20 |
| CORN AND CORN PRODUCTS | 14,870 | 16 |
| COTTON | 6,525 | 7 |
| TOBACCO | 6,186 | 7 |
| RICE | 3,684 | 4 |
| SORGHUM | 3,004 | 3 |
| SUBTOTAL | 71,845 | 79 |
| OTHER LESS ANIMAL PRODUCTS | 19,292 | 21 |
| TOTAL LESS ANIMAL PRODUCTS | 91,137 | 100 |

Percents may not add due to rounding errors.

Rangeland (pasture) and forage (hay) are also important candidates for survey for two reasons: 1) they are complementary or substitute animal foods, and can thus affect the demand for the primary grains; and 2) they act as "confuser" crops in the remote observation of the primary staples.

Table 2-3 indicates that soybeans, wheat and corn head the list of exports by value. They constitute 58% of U.S. staple agriproducts exported during the decade, exclusive of animal products. The other major categories combined - cotton, tobacco, rice, sorghum - represent 21%. The aggregate of the several minor commodities is 21%. No single commodity below the ranking of sorghum produced exports greater than 1% of the total.

Before concluding that the seven crops indicated in Table 2-3 are indeed those of major interest to U.S. agricultural exports, it is well to analyze whether the 10-year cumulation does not smooth out and hide major fluctuations. Figure 2-2 indicates that this is not the case: the relative percentages of the exports of the staple crops vis-a-vis the total agricultural exports have remained relatively constant over the last decade. Within the validity of historical extrapolations, and barring major exogenous disturbances, they can thus be assumed to remain substantially the major U.S. agriexports of the near future.

## 2.5  Derivation of the Countries to be Surveyed

### 2.5.1  Logic

The "universe" for export of U.S. agricultural commodities subdivides as follows:

1.  Countries which do not consume the product at all: i.e., non-customers.

2.  Countries which are consistently self-sufficient (for a given agriproduct which the U.S. produces in surplus): neither customers nor competitors. The sub-category of Countries

# FLUCTUATION OF AGRICULTURAL EXPORTS
## MAJOR AGRICULTURAL EXPORT COMMODITIES AS A FRACTION
## OF TOTAL GROSS AGRICULTURAL EXPORTS, DOLLAR VALUES
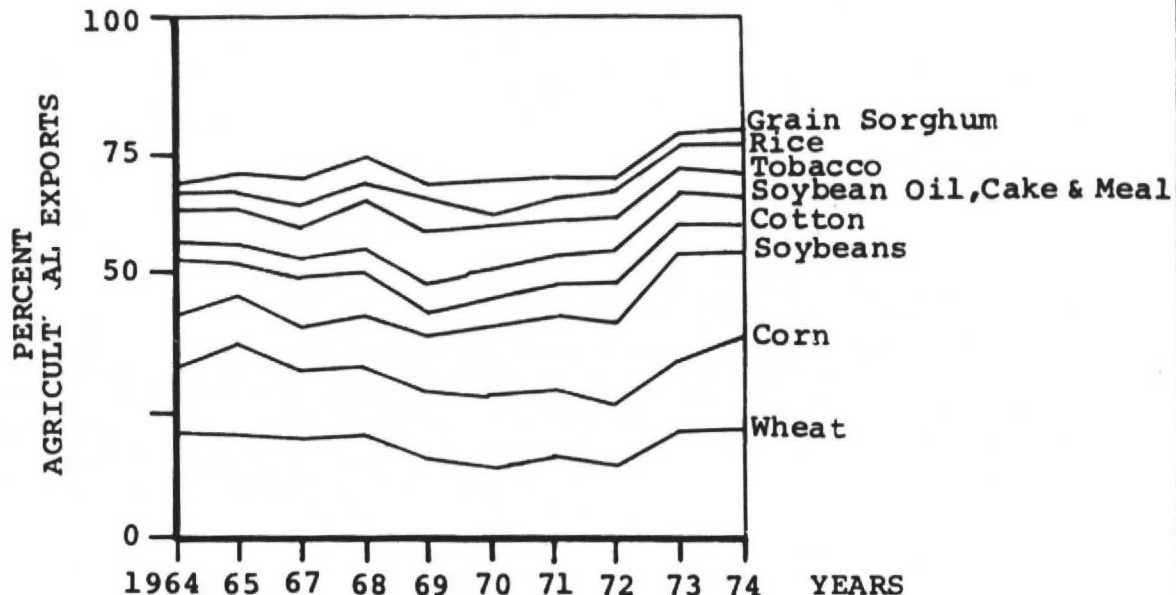


Figure 2-2

# RATIONALE FOR SELECTING COUNTRIES
# FOR AGRICULTURAL SURVEY
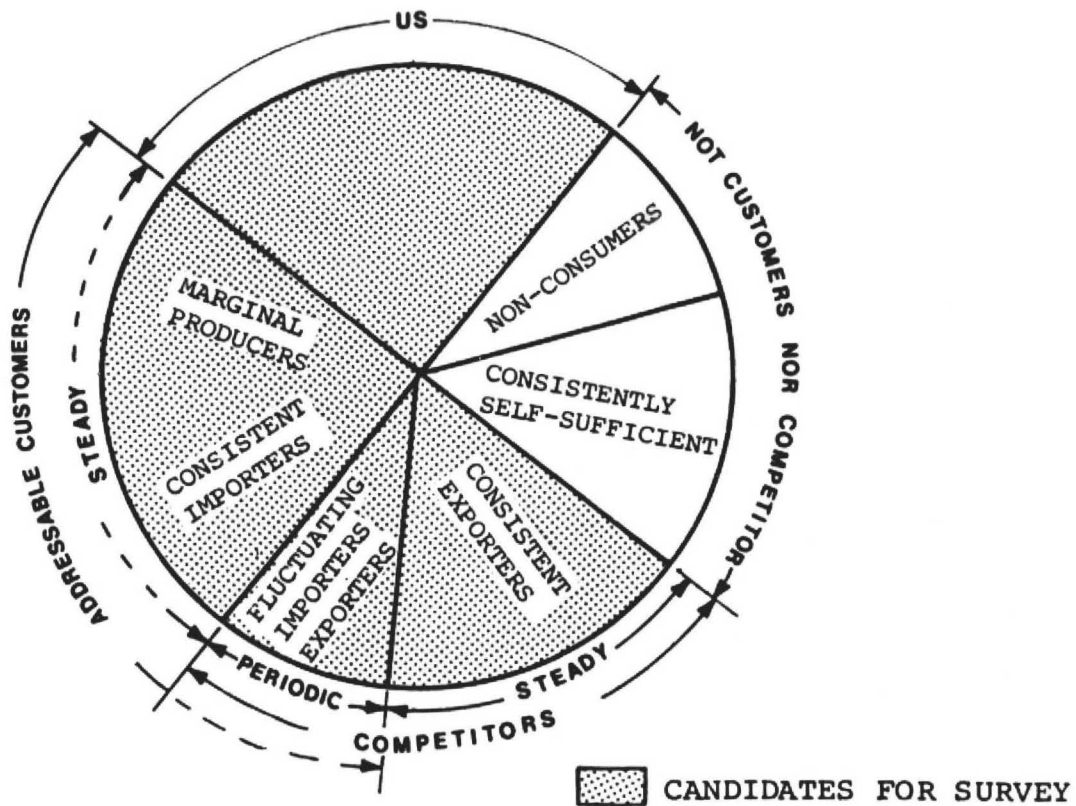


Figure 2-3

which do not purchase from the U.S. for policy reasons, valid
in the 1960's, has, since the massive entry-to-market of USSR,
PRC and Eastern Europe, dwindled into insignificance, at least
for the time being.

3. Countries which produce the product, but not in sufficient amounts
to meet their demand, and which are steady importers.

4. Countries wherein the difference between production and internal
demand alternates between surplus and deficit. These are poten-
tial U.S. customers in shortfall years, potential competitors in
surplus years.

5. Countries which are consistent exporters. These are potential
competitors of the U.S.

6. The U.S. itself, to assess its year-to-year capability to
export.

The requirements for agricultural surveys by the U.S. vary among these
categories of countries.

Countries which exhibit characteristics 1 and 2 - i.e. consistent
non-customers and non-competitors for any given product - are not important
to agricultural U.S. exports of that product. Surveys of their agricul-
tural production of that product are thus of little interest to the U.S.
foreign trade posture.

Surveys of countries in category 3, i.e., which exhibit the qualities
of marginal producers and steady and consistent importers are of interest
to assess the quantity of product they are likely to import in any one year.
This assessment serves as the basis for U.S. agriexport planning on how to

convert this "addressable market" into sales for the U.S. The interest is
proportional to each country's volume of the addressable market and to the
potential competition from other surplus-producing Countries.

Countries of category 4, fluctuating customer-competitors, are of
interest in proportion to the magnitude of the fluctuations of their ag-
ricultural production.

Surveys of countries of Category 5, the potential competitors, are of
interest in proportion to the magnitude of their surplus.

Figure 2-3 depicts the logical choices.

## 2.5.2  Procedure for derivation of the countries to be surveyed

The procedure is to isolate, down to a meaningful level of significance,
the countries which fall in categories 3, 4, and 5, for the principal crops
of interest to U.S. agriexports.

Figure 2-4 depicts the principal customers for the major U.S. export
crops over the last decade. They are ranked from bottom to top by percen-
tage of the dollar volume of imports from the U.S.; from left to right by
total dollar value of each imported commodity. The blank unlabeled areas
at the top of the histogram bars comprise a rather large number of customers,
each of which is significant at most to the 1% level; i.e., its imports
during the decade did not exceed more than 1% of the total U.S. exports of
the corrsponding commodity.

Note that Figure 2-4 indicates percentages and not absolute dollar
values. The dollar purchases of each customer can be obtained by

Figure 2-4

# THE MAJOR CUSTUMERS FOR U.S. AGRIPRODUCTS
## IN PERCENT OF THE CURRENT DOLLAR VALUE OF EACH COMMODITY ACQUIRED
## CUMULATIVE FOR THE DECADE 1964-74.

| WHEAT | SOYBEANS | CORN | COTTON | TOBACCO | SOYBEAN OIL CAKES & MEAL | RICE | GRAIN SORGHUM |
|---|---|---|---|---|---|---|---|
| | | | | | | | BEL-LUX |
| | | BEL-LUX | | | POLAND | UK | NETH. |
| | BEL-LUX | SPAIN | | AUSTRALIA | YUG. | CANADA | |
| | DENMARK | W. GER. | FRANCE | BEL-LUX | BEL-LUX | SAUDI A. | ISRAEL |
| | TAIWAN | | W.GER. | SWEDEN | | W. GER. | |
| YUG. | | CANADA | ITALY | DENMARK | CANADA | JAPAN | |
| W.GER. | CANADA | | INDIA | THAILAND | | S. AFRICA | INDIA |
| PRC | | | HONG KONG | SWITZ. | ITALY | | |
| ALGERIA | | UK | | NETH. | | INDIA | |
| TAIWAN | | | CANADA | | | | |
| UK | W. GER. | | | | NETH. | | |
| NETH. | | ITALY | TAIWAN | JAPAN | | INDONESIA | |
| PAKISTAN | | | | | | | |
| R. KOREA | NETH. | | | | | | JAPAN |
| USSR | | NETH. | R. KOREA | W.GER. | FRANCE | R. KOREA | |
| JAPAN | | | | | | | |
| | JAPAN | | JAPAN | | | | |
| INDIA | | JAPAN | | UK | W. GER. | S. VIET. | |

PRODUCTS ARE RANKED, FROM LEFT TO RIGHT.
CUSTUMERS, FROM BOTTOM TO TOP.

-16-

multiplying the percentages of product acquired, from Figure 2-4, by
the total dollar value of the product exports, from Table 2-3. For
example, Japan's purchases of wheat were 10% of the total U.S. exports;
those of sorghum, 45%. The total U.S. exports were $18.5 Billion for
wheat, $3 Billion for sorghum. Thus Japan's 10% share of U.S. wheat
exports brought a revenue of $1.85 Billion: its purchases of sorghum,
in spite of the much higher share, resulted in a lesser revenue, e.g.
$1.35 Billion.

The information of Figure 2-4 allows elimination of the countries
which fall in categories 1, non-consumers, and 2, self-sufficient. It
aggregates categories 3, net importers, and 4, fluctuating importers-
exporters.

The net importers fall into two sub-categories: steady importers,
typified by Japan, whose purchases essentially repeat year after year;
and fluctuating importers. This second sub-category is clearly of
greater interest to U.S. survey efforts. Figure 2-5 exemplifies its
behavior.

Table 2-4 serves to assess which importing countries fall into
category 4, fluctuating customer-competitors. It indicates that the
USSR is the only such for the commodity wheat, since during the decade
it did export as well as import. The PRC likewise is a fluctuating
customer-competitor for the commodity soybeans, France for corn.

# FIGURE 2-5
## THE MOST FLUCTUATING U.S. CUSTOMERS
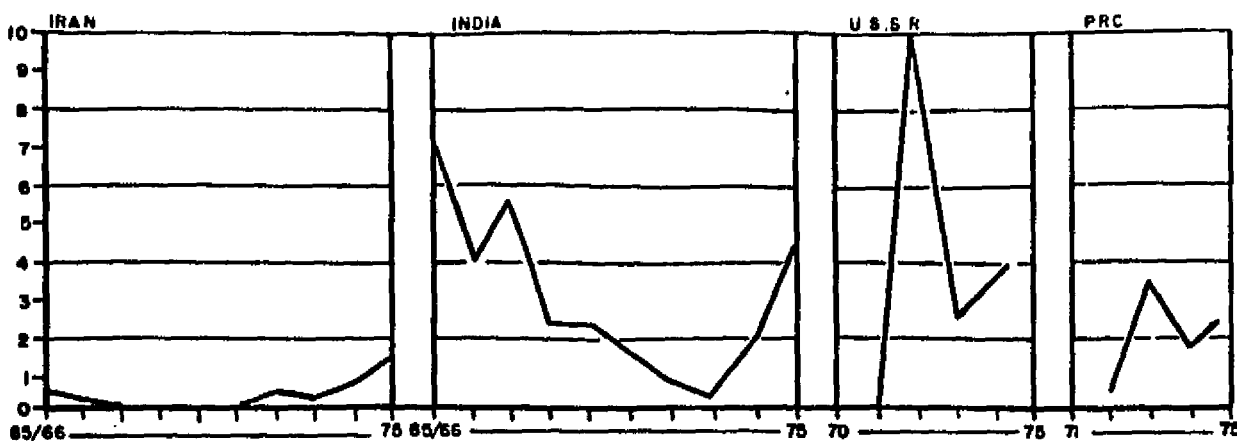


WHEAT
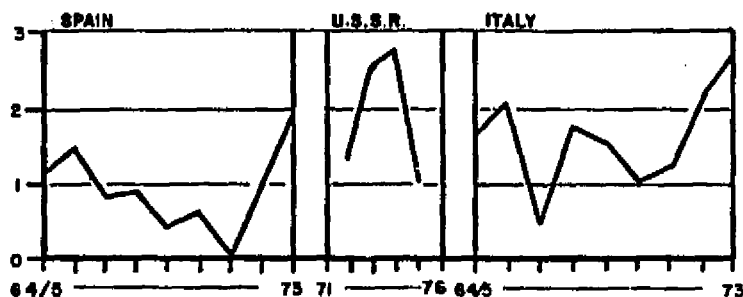


CORN

## THE STEADIEST U.S. CUSTOMERS

YEARS BEGIN
OCT. 1



WHEAT



CORN

MILLION TONS
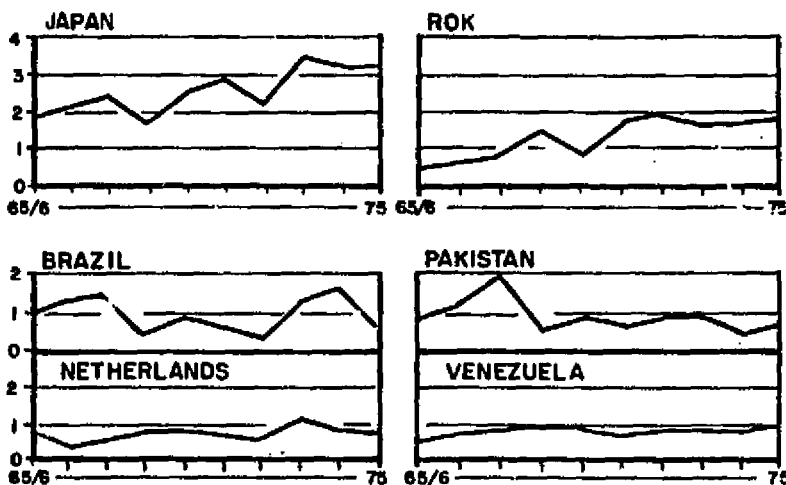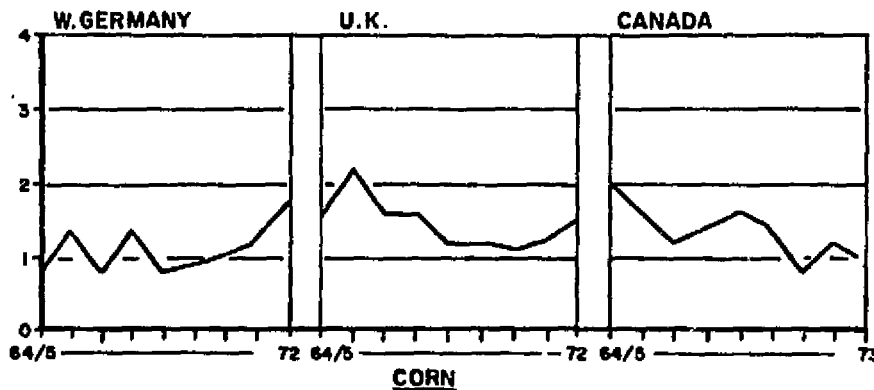
### TABLE 2-4

#### MAJOR WORLD IMPORTS & EXPORTS OF SELECTED AGRICULTURAL COMMODITIES

#### 6-YEAR CUMULATIVE VALUE (1968-73)

#### $M CURRENT - EXPORTS FOB, IMPORTS CIF

| WHEAT | EXPORTS | IMPORTS | CORN | EXPORTS | IMPORTS | SOYBEANS | EXPORTS | IMPORTS |
|---|---|---|---|---|---|---|---|---|
| WORLD | 26,830 | 28,150 | WORLD | 13,600 | 14,930 | WORLD | 9,620 | 9,720 |
| U.S. | 9,740 | - | U.S. | 7,110 | - | U.S. | 8,440 | - |
| CANADA | 5,050 | - | FRANCE | 1,640 | 150 | BRAZIL | 706 | - |
| FRANCE | 3,150 | - | ARGENTINA | 1,620 | - | PRC | 350 | 500 |
| USSR | 2,730 | 2,040 | SOUTH AFRICA | 660 | - | JAPAN | - | 2,590 |
| AUSTRALIA | 2,150 | - | JAPAN | - | 2,540 | FRG | - | 1,620 |
| ARGENTINA | 840 | - | ITALY | - | 1,990 | NETHERLANDS | - | 930 |
| PRC | - | 2,430 | UK | - | 1,380 | SPAIN | - | 910 |
| JAPAN | - | 570 | FRG | - | 1,370 | ITALY | - | 660 |
| UK | - | 2,060 | NETHERLANDS | - | 1,330 | | | |
| INDONESIA | - | 1,500 | SPAIN | - | 970 | | | |
| FRG | - | 1,400 | USSR | - | 630 | | | |
| BRAZIL | - | 1,040 | | | | | | |
| ITALY | - | 840 | | | | | | |
| EGYPT | - | 740 | | | | | | |
| CZECHOSLOVAKIA | - | 610 | | | | | | |
| PAKISTAN | - | 310 | | | | | | |

Table 2-4 also identifies the countries falling within category 5, the steady competitors. These are countries which do not import because they possess a consistent, repeating discretionary surplus: i.e. their production consistently exceeds internal demand.

Figure 2-6 presents a panorama of the principal competitors. It indicates that the U.S. not only possesses the largest but also the most consistent and reliable surplus.

### 2.5.3 Recapitulation of Crops and Countries Candidates for Survey

Table 2-5 recapitulates the principal crops and countries candidates for survey of the major staples for human consumption, down to the 2% level (as percentage of total U.S. agriexports) for the crops,

# Figure 2-6
## DISCRETIONARY SURPLUSES OF MAJOR AGRICULTURAL EXPORTERS

### WHEAT



### AVERAGE (1967-1973)



-20-
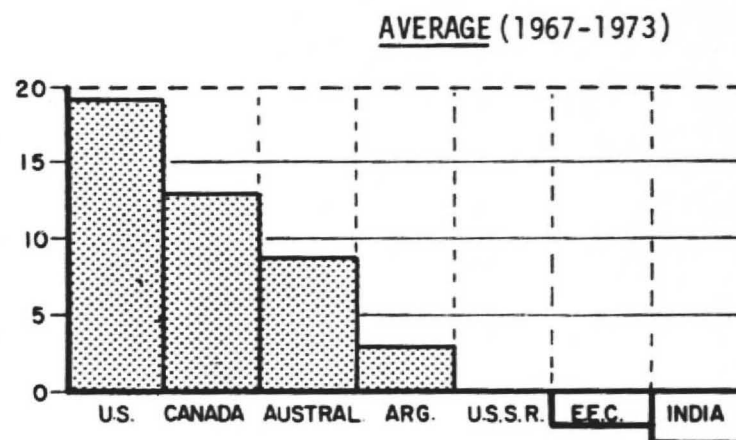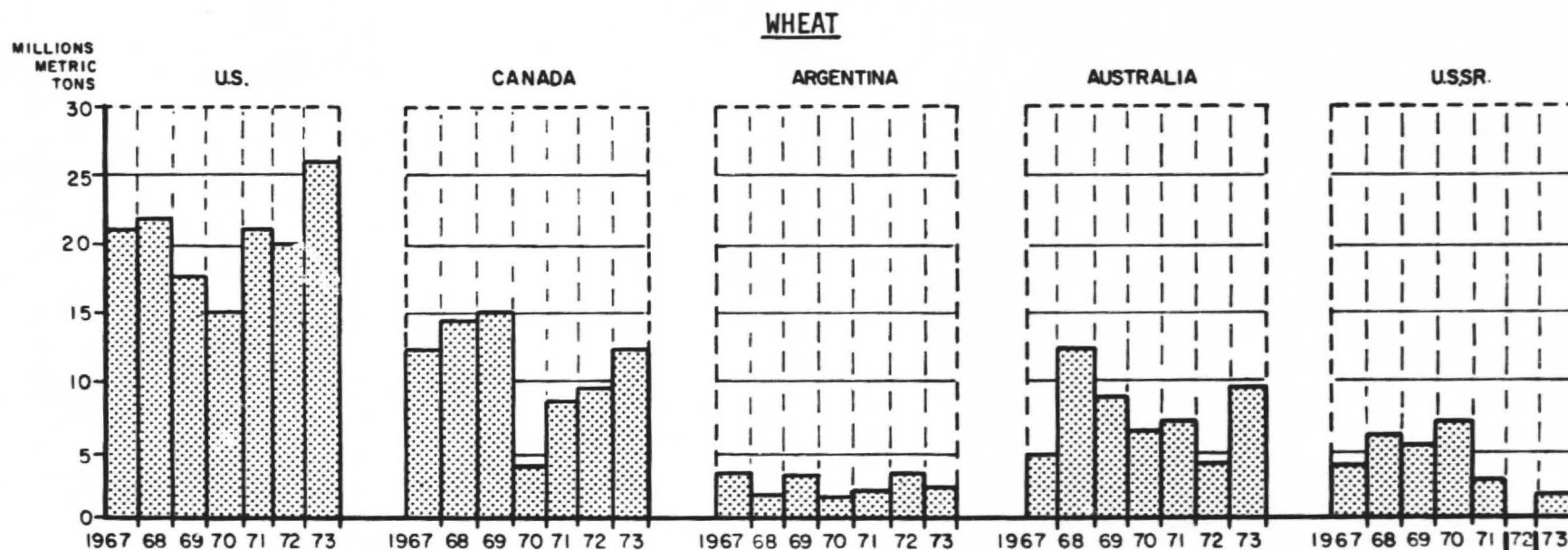
and to the 2% level (fraction of imports of a given commodity from the U.S.) for the Countries. Tobacco is not included since it is not a major staple for human consumption: it is, rather, a luxury item.

TABLE 2-5

CROPS AND COUNTRIES INITIAL CANDIDATES FOR SURVEY

| WHEAT | CORN | SOYBEANS | RICE | COTTON |
|---|---|---|---|---|
| U.S.S.R. | U.S.S.R. | - | - | U.S.S.R. |
| U.S. | U.S. | U.S. | U.S. | U.S. |
| PRC | PRC | PRC | PRC | PRC |
| INDIA | INDIA | INDIA | INDIA | INDIA |
| FRANCE | FRANCE | - | - | - |
| CANADA | - | - | - | - |
| AUSTRALIA | - | - | - | - |
| TURKEY | - | - | - | TURKEY |
| ITALY | ITALY | - | ITALY | - |
| POLAND | - | - | - | - |
| - | BRAZIL | BRAZIL | BRAZIL | BRAZIL |
| - | S. AFRICA | - | BRAZIL | S. AFRICA |
| ROMANIA | ROMANIA | - | - | - |
| - | - | - | INDONESIA | - |
| - | - | - | - | - |
| FRG | - | - | - | - |
| YUGOSLAVIA | YUGOSLAVIA | - | - | - |
| ARGENTINA | ARGENTINA | - | - | ARGENTINA |
| MOROCCO | - | - | - | - |
| - | HUNGARY | - | - | - |
| - | EGYPT | - | - | EGYPT |
| - | - | - | BANGLADESH, JAPAN, THAILAND | - |
| - | - | - | - | PAKISTAN |
| - | - | - | - | MEXICO |

Before concluding that this is indeed the final set of crops and countries, it is important to investigate whether other crops produced by foreign customers are used as substitutes for the primary U.S. exports: for, in this case, and to the extent of the possible substitution, foreign crops other than those indicated in Table 2-5 may become candidates for survey.

## 2.5.4  The Effect of Substitution in Human Nutrition

Past studies and field surveys by U.S. Agencies and Foundations, by
the UN and the FAO, have uniformly concluded that substitution, even
of very similar foods, is not a viable alternative for human consumption
even in periods of famine.  At least, not in the short-term:  habits
and tastes change but slowly [4].

This is confirmed by Figure 2-7, drawn from FAO statistics, which
depicts the historical evolution of human consumption of cereals in
selected countries in the three primary categories of wheat, eaters,
maize eaters, and rice eaters.  Because the per-capita income has grown
steadily and significantly over the years shown, the curves illustrate
the well-known relationship that the consumption of staples grows
initially with increasing per capita income:  beyond a threshold level,
further increases in income result in a decreasing pro-capita consumption
of staples in favor of protein and "luxury" foods.  More significant
to the question of substitution, Figure 2-7 indicates that the ratios of
the diverse staples consumed by a given country change but slowly.  Also,
short term reductions in pro-capita intake of one staple are not
compensated by corresponding increases in other staples.  This indicates
that substitution is not an important driver in human nutrition, except
over relatively long time spans.

FIGURE 2-7

CEREAL SUBSTITUTION IN HUMAN CONSUMPTION
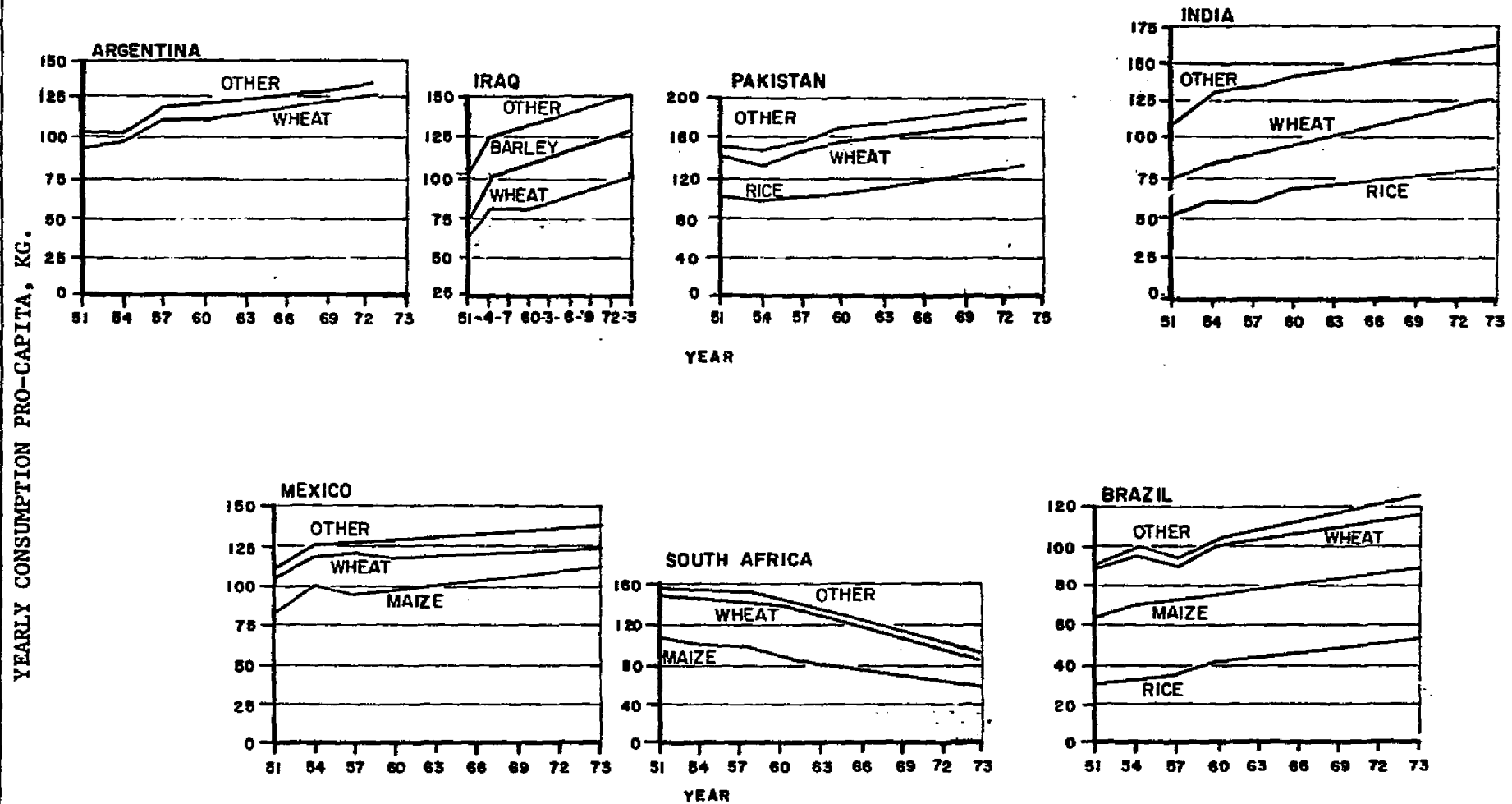
## 2.5.5 The Effect of Substitution in Animal Feed

Over the last fifteen years, the increased demand for protein has caused a significant shift in the composition of animal feed. Appendix A, Figure A-4 illustrates the trend for the USSR and the U.S.: whereas human grain consumption has remained substantially constant, that of feed has significantly increased.

Currently, of the world's cereal harvest only approximately 37% is dedicated to human consumption, the balance 50% after deducting allocations for seed, waste, industrial usage, is devoted to animal consumption.

Contrary to what happens in human consumption, cereals of all types can be substituted in animal feed. As such, their survey is important within the customer and customer-competitor countries already selected in Table 2-5, and in other countries which are major producers of these commodities. The ranking of the major producers of feed grains is presented in Figure 2-8. Cotton is also included.

The complete tabulation of crops and countries which are the final candidates for survey is presented in Table 2-6.

It is well to reiterate that these crops and countries are only those that have passed the criteria established previously. Japan, for example, although a major and steady customer for several US agricultural staples, is not a significant producer of these: as such, it is excluded from the list of candidates for survey, simply because there would not be much to survey.

FIGURE 2-8

RANKING OF WORLD PRODUCERS OF PRINCIPAL CEREAL SUBSTITUTION AND FIBER CROPS

SOURCE: USDA AGRICULTURAL
STATISTICS

% WORLD TOTAL

TOTAL 155 MILLION METRIC TONS

U.S.S.R
WEST EUROPE
DENM
W.GER.
NORTH AMERICA
ASIA
UK
USA.
INDIA
EAST EUROPE
TURKEY
FRANCE
CANADA
PRC
E. GER
CZECH
POLAND
AFRICA
OCEANIA
AUSTRA
S. AMERICA

BARLEY

% WORLD TOTAL

TOTAL 54 MILLION METRIC TONS

U.S.S.R
NORTH AMERICA
CANADA
W.EUROPE
SWEDEN
FRANCE
E.EUROPE
U.S.A.
E.GER.
CZECH
W.GER.
POLAND
ASIA
OCEANIA
PRC
AUSTRA.
S.AMERICA
ARGENT.

OATS

% WORLD TOTAL

TOTAL 29 MILLION METRIC TONS

E.EUROPE
U.S.S.R.
E.GER,
POLAND
W.EUROPE
AUSTRIA
W.GER.
N.AMERICA
CANADA
S.AMERICA
ASIA
U.S.A.
ARGENT.
TURKEY

RYE

% WORLD TOTAL

TOTAL 14 MILLION METRIC TONS

ASIA
TURKEY
N.AMERICA
PAKIST.
MEXICO
U.S.S.R.
INDIA
AFRICA
S.AMERICA
U.S.A.
ARGENT.
COLUMBIA
PRC
S.AFRICA
EGYPT
BRAZIL
W.EUROPE
GREECE

COTTON

**TABLE 2-6**

**CROPS AND COUNTRIES FINAL CANDIDATES FOR SURVEY**

| WHEAT | CORN | SOYBEANS | RICE | BARLEY | OATS | RYE | COTTON | SORGHUM |
|---|---|---|---|---|---|---|---|---|
| U.S.S.R. | U.S.S.R. | - | - | U.S.S.R. | U.S.S.R. | U.S.S.R. | U.S.S.R. | - |
| U.S. | U.S. | U.S. | U.S. | U.S. | U.S. | U.S. | U.S. | - |
| PRC | PRC | PRC | PRC | PRC | PRC | - | PRC | - |
| INDIA | INDIA | - | INDIA | INDIA | - | - | INDIA | INDIA |
| FRANCE | FRANCE | - | - | FRANCE | FRANCE | - | - | - |
| CANADA | - | - | - | CANADA | CANADA | - | - | - |
| AUSTRALIA | - | - | - | AUSTRALIA | AUSTRALIA | - | - | - |
| TURKEY | - | - | - | TURKEY | - | TURKEY | TURKEY | - |
| ITALY | ITALY | - | ITALY | - | - | - | - | - |
| POLAND | - | - | - | POLAND | POLAND | POLAND | - | - |
| - | BRAZIL | BRAZIL | BRAZIL | - | - | - | BRAZIL | - |
| - | S. AFRICA | - | - | - | - | - | S. AFRICA | - |
| ROMANIA | ROMANIA | - | - | - | - | - | - | - |
| - | - | - | INDONESIA | - | - | - | - | - |
| - | - | - | - | - | - | - | - | - |
| - | - | - | - | UK | - | - | - | - |
| FRG | - | - | - | FRG | FRG | FRG | - | - |
| YUGOSLAVIA | YUGOSLAVIA | - | - | - | - | - | - | - |
| ARGENTINA | ARGENTINA | - | - | ARGENTINA | ARGENTINA | ARGENTINA | ARGENTINA | - |
| MOROCCO | - | - | - | MOROCCO | - | - | - | - |
| - | HUNGARY | - | - | - | - | - | - | - |
| - | EGYPT | - | - | - | - | - | EGYPT | - |
| - | - | - | BANGLADESH, JAPAN, THAILAND | - | - | - | - | - |
| - | - | - | - | DENMARK | - | - | - | - |
| - | - | - | - | CZECHOSLOVAKIA | CZECHOSLOVAKIA | - | - | - |
| - | - | - | - | GDR | GDR | GDR | - | - |
| - | - | - | - | - | SWEDEN | - | - | - |
| - | - | - | - | - | AUSTRIA | - | - | - |
| - | - | - | - | - | - | - | PAKISTAN | - |
| - | - | - | - | - | - | - | MEXICO | - |

## 2.6   Effect of Global Crop Forecast on Other U.S. Goals

### 2.6.1   Effect of the requirement to maximize consumer satisfaction

In a free-market economy, early prediction of the status of crops
will result in the maximum possible reduction of U.S. consumer prices
for the primary commodities.  The U.S. spends approximate one-half of
its revenues from the exports of agriproducts in importing complementary
and suppementary agricultural commodities and derived products.
Complementary imports are items which are not produced in the U.S.: e.g.
coffee, bananas, tea.  Supplementary imports consist of products
available domestically, but which consumers cherish for reasons of taste or
custom:  e.g. wines, canned tomatoes, pasta products.

The complementary and supplementary agricultural products play a
significant role in insuring the satisfaction of U.S. consumers.  Thus
inventories of the respective producing countries - for example,
Brazilian coffee - are important, and a scenario analogous to, but in
reverse of that presented above for the exports, could be constructed.
This was, however, beyond the scope of this study.

### 2.6.2   Role of Global Crop Forecast as Tool of Foreign Policy

The optimum trading of U.S. -produced food in exchange for political
advantages requires the early knowledge of the requirements of the numerous
countries who are actual or potential AID recipients.  This implies that
the survey of principal crops should be extended to at least the most im-
portant of these countries.  The rationale for selecting target crops
and countries differs from that developed above:  it addresses the selection
objectives of U.S. Foreign Policy.  Crops different from those thus far
identified do emerge -- e.g. cassava.

This investigation was beyond the scope of this study.

## 2.7  The International Scenario

The contribution achievable to the UN goal of "world peace" by providing adequate nutrition to the world's masses is a matter of philosophical speculation.  History provides abundant examples of wars and of local disturbances induced by want; as well as examples of prolonged mass deprivations not leading to such violent epilogues; and also instances of major wars initiated by well-fed contenders.

The second UN goal, "promoting global freedom from famine" is more specific and amenable to quantitative analysis, without the need to pass judgements on its philosophical correctness.  It is presented in Appendix A.

The UN/FAO scenario was not utilized for this study.  It is however worth noting, as shown in Appendix A, that this scenario maintains the crops and countries of the U.S. Balance of Trade scenario, except for the inclusion of additional countries producers and consumers of the staple crop rice.

## 3.0 SYSTEM PERFORMANCE REQUIREMENTS AND CURRENT PERFORMANCE

### 3.1 Derivation of the Data System's Performance Requirements

Appendix B presents an overview of the agricultural survey methods currently in use throughout the world. As indicated therein, crop production forecasts comprise two distinct functions:

- Measurement of the current status of the crop. For reasons of cost, this measurement is always performed by statistical sampling. Following the USDA designation, it is also called "estimate."

- Extrapolation to the future harvest, more properly termed "forecast." This function is accomplished based on historical regressions augmented by "best estimates" of the future incidence of adverse or beneficial phenomena affecting the crop, such as weather.

The estimate of the current status is composed of two subfunctions:

- Estimate of the acreage under cultivation

- Estimate of the crop's condition.

This study focuses upon the acreage estimates, because they are currently amenable to measurement from orbital data systems.

Table 3-1 presents the accuracies of acreage mensuration achieved by various countries and by U.S. States and Counties [5],[6],[7]. It is subject to the following interpretation:

● The reported accuracies were obtained from direct discussions with
responsible individuals in the US and abroad, augmented by analysis
of their data.

● The reported accuracies are valid late in the season; i.e. not
earlier than at least one month prior to harvest.

● The stated accuracies refer to sampling errors only, i.e. the com-
bination of the random errors induced by the limited number of
samples and by the accuracies of measuring each sample's crop
acreage. Systematic errors are not included.

TABLE 3-1

ACCURACIES CURRENTLY ACHIEVED IN THE MEASUREMENT OF ACREAGE

| COUNTRY | CROP | ACREAGE MEASUREMENT ERROR PREHARVEST ONE SIGMA |
|---|---|---|
| U.S. | WHEAT | ±2% |
| | CORN | ±6% |
| | SOY | ±6% |
| ITALY | WHEAT | ±4% |
| | FORAGE | ±8% |
| | CORN | ±7% |
| WESTERN EUROPE | SIMILAR TO ITALY | - |
| CANADA | SPRING WHEAT | ±3% |
| USDA-FAS WORLD ESTIMATE | WHEAT | ~±3% |
| U.S. STATES (without auxiliary list sampling) | MAJOR STATE CROP | ~±7%-15% |
| U.S. COUNTIES (without auxiliary list sampling | MAJOR COUNTY CROP | UP TO ±20-25% |

Perusal of Table 3-1 leads naturally to the question as to whether the accuracy figures therein presented do indeed represent the user requirements. In numerous discussion by ECOsystems personnel with domestic and foreign user agencies, the following transpired:

- Currently achieved accuracies *obtained at the current statutory reporting dates* cannot be reduced.

- They must be achieved at costs not in excess of current costs: in fact, it is desirable that the costs be lower.

- Accuracies lower than those currently achieved are only useful if they can provide valuable adjuncts to the current conventional systems, either in terms of significantly lowering costs or of providing significantly improved performance of the total system.

- Users are willing to pay a modest premium for significantly improved accuracies achieved *at the statutory reporting dates*.

- Users are disposed to compromise accuracy for timeliness, i.e. less accurate results obtained earlier than currently achieved are acceptable.

- The combination of accuracy and timeliness warrants a significantly higher premium than what users are willing to pay to accuracy alone.

The current "user market" situation strongly counsels that the accuracy levels depicted in Table 3-1 be assumed as goals for the data system's performance.

3.2  Assessment of the Performance of Orbital Data Systems in Estimating
     Acreage

The purpose of this investigation was to quantify numerically the
performance of existing orbital Data Systems, for comparison against the
user requirements depicted in Table 3-1.  The investigation was performed
in two steps:

- First, analysis of all available reports on crop acreage mensu-
  ration by ERTS investigators:  these represent tests performed
  during a concentrated activity in the early phases of the LANDSAT
  program, from its launch in July of 1972 until 1974.

- Second, analysis of findings from later tests, to assess the
  trend of improvement.

To this effect, results were gathered from available, published ERTS
investigation reports containing numerical performance figures.  Details
of the analysis are presented in Appendix C.

The resulting statistical population consisted of 224 distinct tests,
performed in 10 locations, embracing 7 crops.

The cumulative curve of results is presented in Figure 3-1A.  It
graphs the accuracies achieved versus the percentage of the tests achieving
them.  It presents the results for both the inventory and the mapping, or
land use, mode.  Inventory mode means the determination of the proportion
of a given crop within the area under examination, without however speci-
fying the crop's geographic location.  It is the mode universally used for
estimating agricultural acreage.  Mapping mode, as the word indicates,
connotes the identification not only of the species sought, but also of

# FIGURE 3-1

## DISTRIBUTION OF CLASSIFICATION ACCURACIES FOR ERTS INVESTIGATIONS ON ACREAGE MENSURATION

### A. CLASSIFICATION ACCURACY OF ERTS INVESTIGATIONS
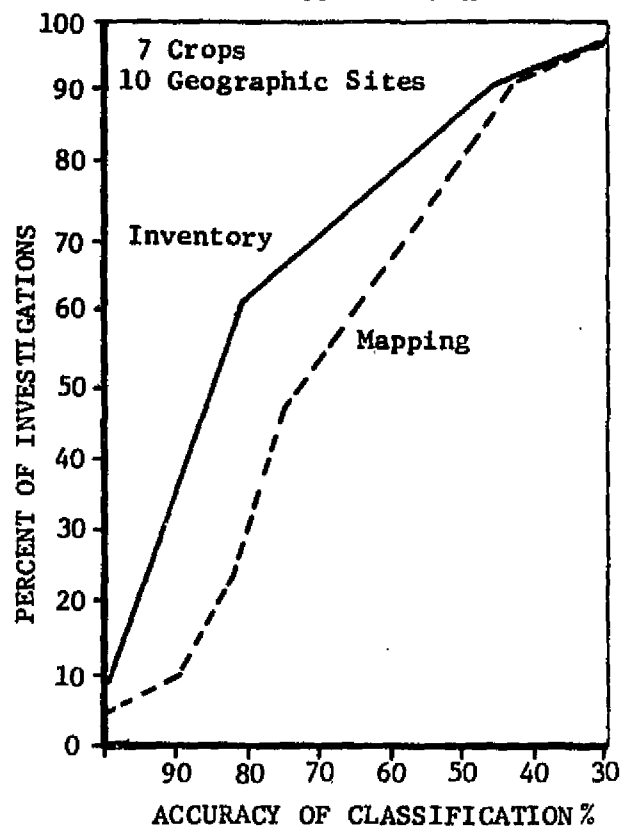


NUMBER OF DATA POINTS: 224

### B. CLASSIFICATION ACCURACY OF ERTS INVESTIGATIONS WITH PROPORTION OF TARGET CROP >10%
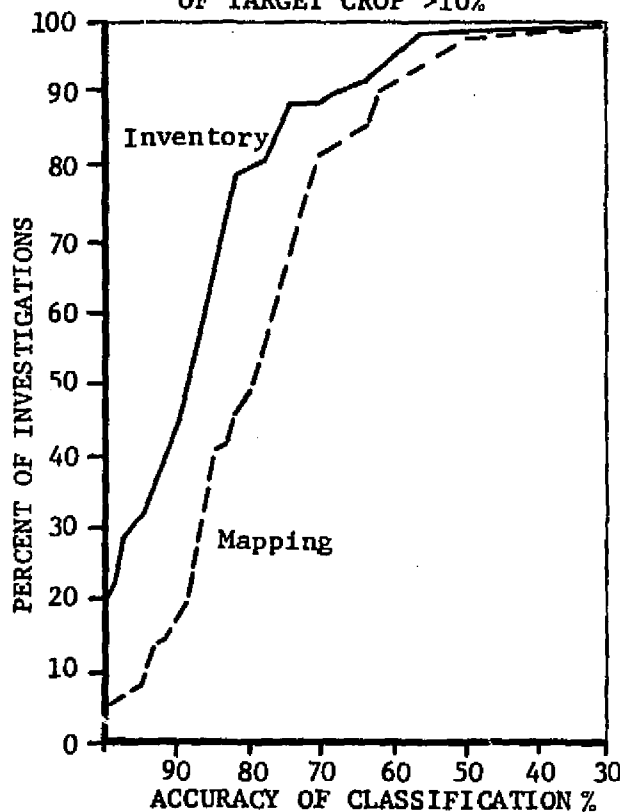


NUMBER OF DATA POINTS: 64

their geometric boundaries and geographic locations. This mode is generally not employed in conventional agricultura. acreage estimation, due to its higher cost with respect to the inventory mode. It is presented in this report because of its implications to the theory of discrimination discussed in Chapter 5.

It can be noted from Figure 3-1 that the accuracies achieved by ERTS investigators in the mapping mode are lower than those competing to the inventory mode: this is quantitatively consistent with the fact that the mapping mode requires more information than proportion estimation: conversely, for equal information content, its level of performance should be lower.

Since, as will be shown in Chapter 5, the accuracy of multispectral classification worsens when the proportion of target crops to the total scene is small, a subset was extracted from the total statistical population, based on the criterion of 10% or larger pr.portion of target crops.

This yielded 64 distinct data points: the test results are presented in Figure 3-1B. Note the higher accuracies achieved.

The probability density functions corresponding to the curves of Figure 3-1 are shown in Appendix C. Computation of the accuracies as a function of confidence limits from these yields the classification results depicted in Table 3-2. It shows that the one sigma accuracy of classification in the *inventory mode* for the total population of 224 data points is 74%: this increases to 86% for the subset of 64 data points featuring 10% or greater proportions.

TABLE 3-2

**ACCURACY OF CLASSIFICATION OF CROPS ACHIEVED**

**BY ERTS INVESTIGATORS**

Universe of 224 classification tests on 7 crops, 10 geographic locations, one sigma confidence

|  | PROPORTION ESTIMATION | MAPPING |
|---|---|---|
| ACCURACY % | 74 | 63 |
| ERROR % | 26 | 37 |

64 classification tests extracted from the above universe, wherein the crop-to-total-area proportion was greater than 10%.

| ACCURACY % | 86 | 76 |
|---|---|---|
| ERROR % | 14 | 24 |

Comparison with Table 3-2 shows that the average performance of ERTS investigator tests in accomplishing inventory fell short of the accuracies currently achieved by conventional methods, with the possible exception of inventories performed at county levels.

ERTS investigations ceased after 1974. To assess whether the early performance has improved in the approximately four years elapsed since, a number of subsequent investigations, performed by reputable organizations, were analyzed. The details are presented in Appendix C: the overall results are shown in Table 3-3. It can be seen that there has been improvement since 1974: yet performance is still short of the user's requirements.

TABLE 3-3

ACCURACY OF CLASSIFICATION OF CROPS ACHIEVED FROM 1974 TO 1978

Universe of 18 classification endeavors on 10 crops, 13 geographic
locations; one sigma confidence.

|  | PROPORTION ESTIMATION | MAPPING |
|---|---|---|
| ACCURACY % | 86 | 73 |
| ERROR % | 14 | 27 |

The logical next inquiry is to ascertain whether the 75-85% inven-
tory accuracy thus far achieved by LANDSAT experimenters is a limit im-
posed by nature, or whether there exist driving elements in the Data
System whose exploitation can further enhance crop discrimination and
identification.

## 4.0 INVESTIGATION INTO THE PHYSICAL REASONS FOR THE DISCREPANCY BETWEEN REQUIREMENTS AND PERFORMANCE

### 4.1 Methodology

The generalized end-to-end data system of radiometric earth-observation platforms employed currently and expected to be used in the 1980-1990 decade is depicted in Figure 4-1. The methodical investigation of its performance was begun at the front-end, by analyzing the relationships between the information provided by nature, i.e. the crop spectra, and the measurement capabilities of radiometric sensing.

The initial question addressed was whether reflectance spectra of crops do provide sufficient information for accurate identification or discrimination, and, if so, under what operating conditions. By discrimination is meant the recognition that the observed classes are distinct: identification, in addition, is the attribution of a specific label, e.g. agricultural crop, to each distinct class. The difference between discrimination and identification is conceptually akin to that between inventory and mapping.

Proper performance of this investigation suggests the removal, insofar as possible, of all effects which perturb the spectral reflectance patterns: primarily atmospheric absorption, path radiance, instrumental drifts, effects of varying orientations of illumination and look angles. This requires the availability of "pure" spectra, i.e. of fine-grained crop spectral measurements taken under carefully calibrated conditions.

FIGURE 4-1

FUNCTIONAL DATA FLOW GENERAL RADIOMETRIC OBSERVATION SYSTEM

ON-BOARD

$$n = K \sqrt{\Delta\lambda \cdot \rho} \; r^2$$

STORAGE

COMPRESSION PROCESSING

SENSOR → FORMATTING → ⊘ → ⊙ → MODULATOR TRANS-MITTER

E/M LINK

DATA RELAY → A

$$q = k \frac{A}{r^2 \tau} \; \ell n_2 \; n$$

GROUND RECEPTION AND PREPROCESSING

A → RECEIVER FRONT-END → VIDEO PROCESSOR → HD STORAGE → CORRECTION → B

B → CCT's

B → HARD COPY RECORDER → IMAGES

USER PROCESSING

DATA PRODUCTS

→ ARCHIVE

→ DISSEMINATION

→ PRE-PROCESSING → INFORMATION EXTRACTION → USER PRODUCT

n = NO. RADIOMETRIC LEVELS
K = SYSTEMS CONSTANT
$\Delta\lambda$ = WAVELENGTH INTERVAL
$\rho$ = REFLECTANCE OF IMAGED ELEMENT
r = GROUND RESOLUTION

q = DATA RATE, BITS/SEC.
A = AREA OF SCENE
$\tau$ = TIME OF IMAGING SCENE
k = COMPRESSION RATIO

―― CURRENT SUBSYSTEMS
―― POSSIBLE FUTURE SUBSYSTEMS

-38-

Several spectral data bases were available in the literature prior
to 1975.  Table 4-1 lists the principal ones.

TABLE 4-1

PRINCIPAL SPECTRAL DATA BASES AVAILABLE IN THE LITERATURE

| DESIGNATION | SPONSOR AGENCY | PUBLICATION DATE | LOCATION |
|---|---|---|---|
| MTF | NASA/MTF | 12/73-5/74 | MTF - BAY ST. LOUIS, MS |
| LARS | NASA | 7/68 | LARS - WEST LAFAYETTE, IN |
| ERIM | U.S. AIR FORCE | 1972 | NTIS - SPRINGFIELD, VA |
| ERDL | U.S. ARMY | 1961-1962 | NTIS - SPRINGFIELD, VA |
| WESLACO | USDA/ARS | 10/66-9/67 | USDA ARS - WESLACO, TX |
| KRINOV | USSR | 1936-1938 | NRC - OTTAWA, CANADA |
| IITRI* | NASA/USAF | VARIOUS | NTIS - SPRINGFIELD, VA |

*Summary of data from other sources.

Unfortunately, most of these spectra do not correspond to ideal condi-
tions, because:  1)  they are either spectra of plant leaves taken in the
laboratory, thus not of the plants as a whole; 2)  calibrations surrounding
the taking of whole-plant spectra were not consistent.  Nevertheless,
these spectral data bases were useful as secondary, qualitative sources of
information.

The first set of "pure" spectra became available from Bushland, Texas
in 1975.  They were supplied to ECOsystems through the courtesy of Dr.
James C. Harlan of the Spectral Characteristics Laboratory, Remote Sensing
Center, Texas A&M University.  Dr. Harlan was also most generous with minu-
tious information as to the experimental procedure followed, ground truth

maps, files and aerial imagery. This allowed certification of the quality of these spectra and permitted the compensation of spurious effects which had crept into the data gathering procedure in spite of the experimenter's diligence.

Subsequently, other data sets, also collected under calibrated conditions at low altitude, became available. The set of data used in this study is indicated in Table 4-2.

## TABLE 4-2

### CHARACTERISTICS OF THE "PURE" SPECTRAL DATA USED IN THE ANALYSIS

|  | BUSHLAND, TX | HAND CO., SD | IMPERIAL VALLEY, CA |
|---|---|---|---|
| TYPES OF CROPS | W, R, P, (C,S) | W, A, P, O, B (C) | (W, S, SB, AS, A, ME) |
| NUMBER OF FIELDS | 100 | ≈15 | ≈100 |
| SPECTRAL RANGE, MICRONS | 0.4 to 2.4 | 0.4 to 2.4 | 0.4 to 1.0 |
| SPECTRAL RESOLUTION, MICRONS | 0.02 from 0.4 to 1.1 | 0.02 from 0.4 to 1.1 | 0.0014 |
| NUMBER OF SPECTRA GATHERED PER FIELD | ≈20 | ≈20 | ≈15 |
| CALIBRATION PROCEDURE | OVERFLIGHT OF REFERENCE PANEL @ 20' INTERVALS | | REFERENCE SOURCE BEFORE & AFTER FLIGHT |
| INSTRUMENT | S-191H CONICAL WHEEL SPECTROMETER | | COLLINS SPECTROMETER |
| PLATFORM | HELICOPTER | HELICOPTER | AIRCRAFT |
| DATES OF GATHERING | 1974-75 CROP SEASON | 1976-77 CROP SEASON | 1975 GROWING SEASON |
| GROUND RESOLUTION, METERS | 23.8 | 23.8 | 18 X 18 |
| FLIGHT ALTITUDE, METERS | 60.9 | 60.9 | 610 |

W = WHEAT    R = RYE    P = PASTURE    C = CORN    S = SORGHUM    A = ALFALFA    O = OATS    B = BARLEY    SB = SUGAR BEETS

AS = ASPARAGUS    ME = MELONS    (  ) = SPECTRA AVAILABLE DURING PARTIAL GROWTH CYCLE ONLY.

These data sets represent a major improvement with respect to the earlier spectra available in the literature. Nevertheless, they still present significant uncertainties. For example: none of the data sets embraces a temporal span sufficiently wide so as to contain all the phenologic stages of corn or sorghum. The data-gatherer's primary

interest appears to have been the measure of wheat. In several cases, obvious sensor and/or processing errors have crept in. Frequently, the percent of crop ground cover is not unambiguously stated: thus several crop spectra are mixed with soil spectra in such fashion that it is difficult to unravel the true reflectance of either. In many cases, the nomenclature does not correspond with the actual ground cover: fields planted to a certain crop which never does emerge are labeled as containing that crop. No attempt is apparent to reconcile widely disparate reflectances pertaining to the same field, as for example the reflectances of soil, labeled "bare," before emergence and after harvest.

This state of affairs, particularly evident in the Hand County spectra, requires considerable judgement, and inter-comparison of spectra from all the data sets, in order to avoid the conclusion that crop reflectance spectra are essentially random.

## 4.2 Results from the Analysis of the Field Spectra

The details of the analysis of the available data sets are presented in Appendix D. The salient conclusions are:

1. The shapes of the crop reflectance spectra are significantly similar at similar growth stages - i.e. after the crops are sufficiently developed to cover the underlying soil. As exemplified in Figure 4-2, there is little evidence of sharply unique spectral structures among the various crops.

2. The spectral shapes remain significantly similar throughout the crop's growth stages, provided that the crop's ground cover is sufficiently dense so as to minimize interference from soil reflection.

# FIGURE 4-2
## SPECTRA OF VARIOUS CROPS AT 100% LEAF COVER IN IMPERIAL VALLEY, CALIFORNIA
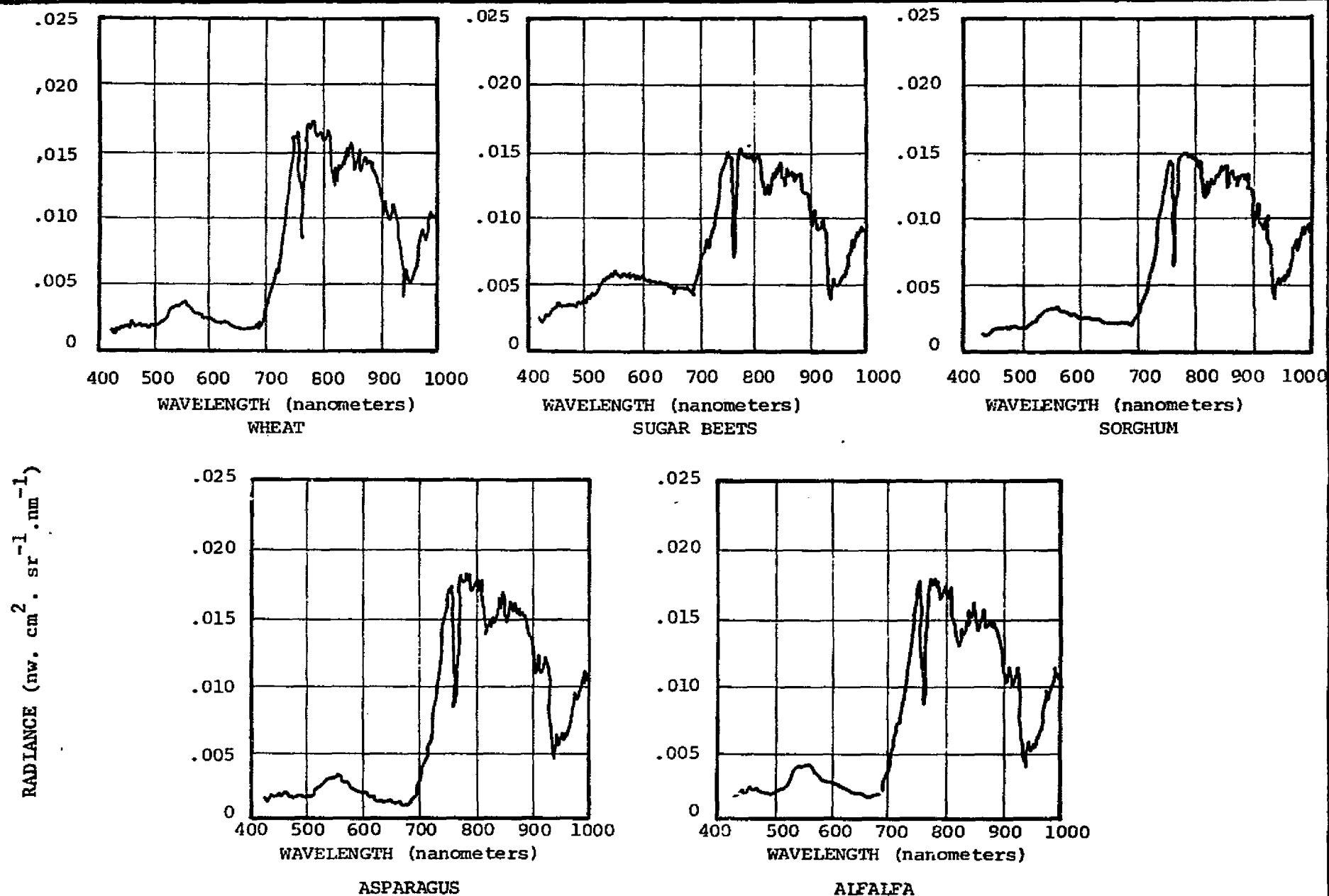


WAVELENGTH (nanometers)
WHEAT

WAVELENGTH (nanometers)
SUGAR BEETS

WAVELENGTH (nanometers)
SORGHUM

RADIANCE (nw. cm$^2$ . sr$^{-1}$ .nm$^{-1}$)

WAVELENGTH (nanometers)
ASPARAGUS

WAVELENGTH (nanometers)
ALFALFA

SOURCE: Atlas of Selected Crop Spectra Imperial Valley, California
S.G. Ungar, NASA Institute for Space Studies, 1977

-42-

3. The spectra exhibit significant differences of mean amplitudes
   as a function of stage of growth. Typicaly progressions are
   exemplified in Figure 4-3.

4. The mean values of the spectral amplitudes exhibit fluctuations.
   This is exemplified in Figure 4-4. Fluctuations are present
   within single fields and in ensembles of fields, and occur in
   all regions of the spectrum.

5. The fluctuations among ensembles of fields are generally larger
   than those pertaining to single fields. A possible explanation
   of the intrafield fluctuations is the presence of discontinui-
   ties -- patches, bare spots, stressed areas -- within the fields;
   and the clutter effect originating from the varying orientations
   and sizes of the reflecting components of the canopy. A logically
   appealing explanation of the interfield fluctuations is that
   cropped fields do not reach the same development stage simul-
   taneously. Figure 4-5 depicts for various crops the fraction,
   as a function of time, of the cropped surface which reaches a
   given stage of phenologic development. It can be seen that as
   many as 60 days are required to progress from 0% to 100% of the
   covered area. Thus an ensemble of fields generally contains in-
   dividual fields at different phenologic development, and hence
   different reflectances. The presence of stress causes similar
   effects. The available data, cf. Appendix D, suggest 0.1 as a
   typical lower value and 0.3 as a typical upper bound of the coef-
   ficients of variation for an ensemble of fields.

6. Identification was not found possible from spectral signature
   or from differences in amplitude alone. It requires additional
   external information, e.g. knowledge of the phenologic progres-
   sions prevailing within the region being observed.

# FIGURE 4-3

TEMPORAL PROGRESSION OF THE MEAN REFLECTANCE FOR THE CROPS CONTAINED

IN THE HAND COUNTY DATA SET

REFLECTANCES ARE ADJUSTED TO 100% GROUND COVER



SPRING WHEAT, FIELD 169 —··—    OATS, FIELD 136 —··—    BARLEY, MULTIFIELD —·—

PASTURE, FIELD 198 ······    ALFALFA, MULTIFIELD ——

-44-
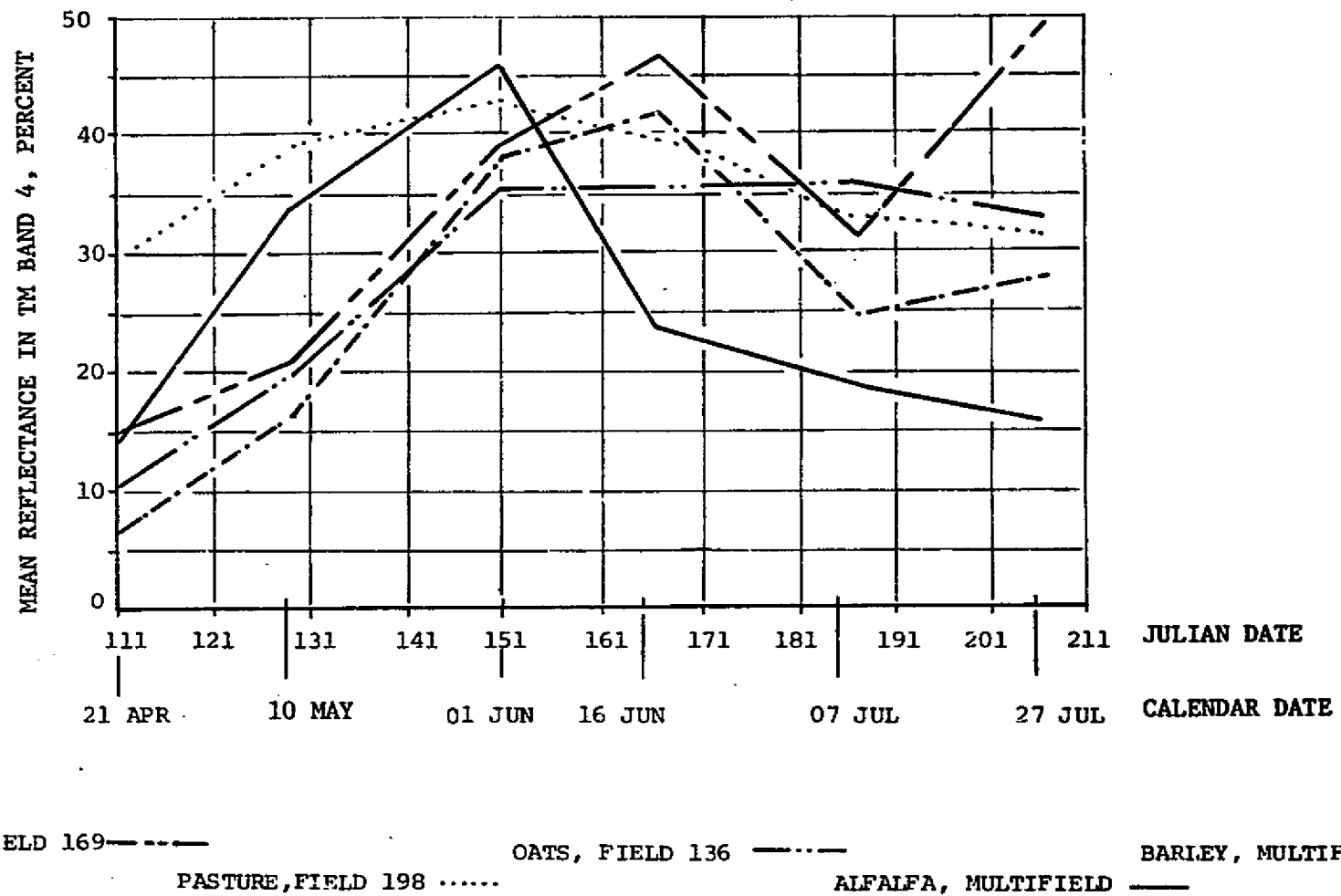
FIGURE 4-4

SPECTRAL VARIANCES OF CORN

HAND COUNTY, N.D., 27 JULY 1977



FIELD NO. 76

ALL CORN FIELDS
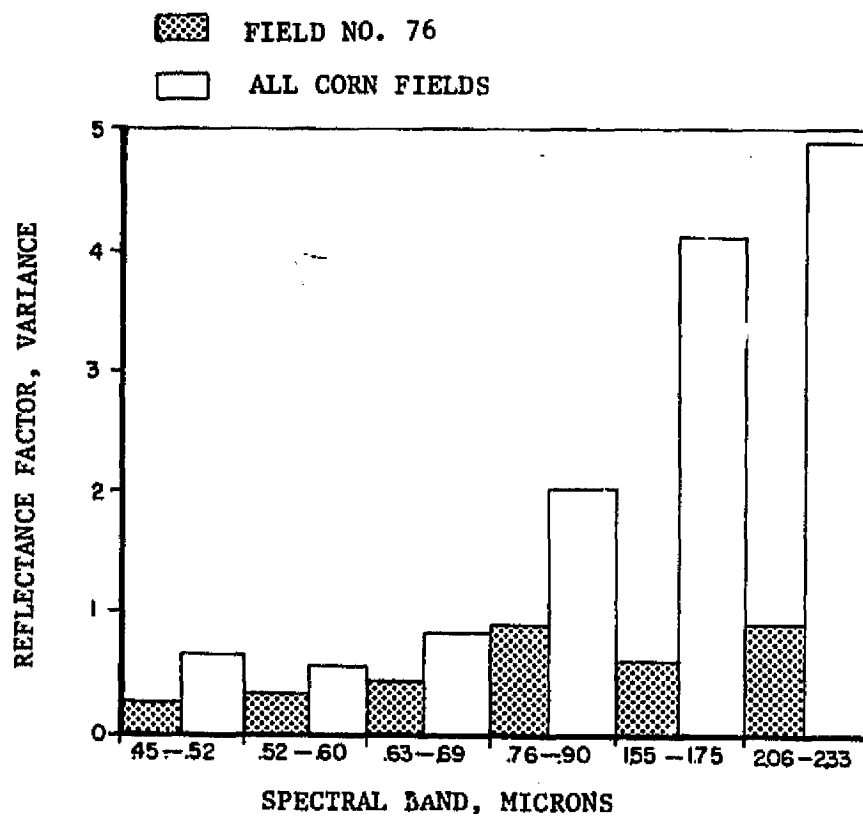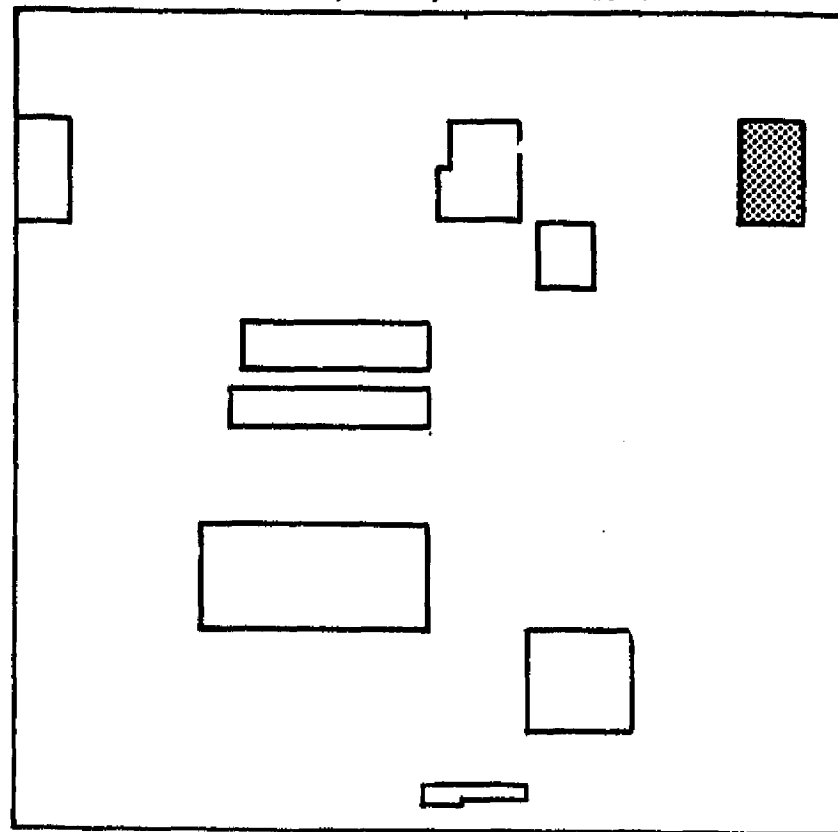
# FIGURE 4-5

## AREAL DEVELOPMENT OF CROPS AS A FUNCTION OF TIME SOUTHWEST CROP
## REPORTING DISTRICT, KANSAS, 1973 SEASON
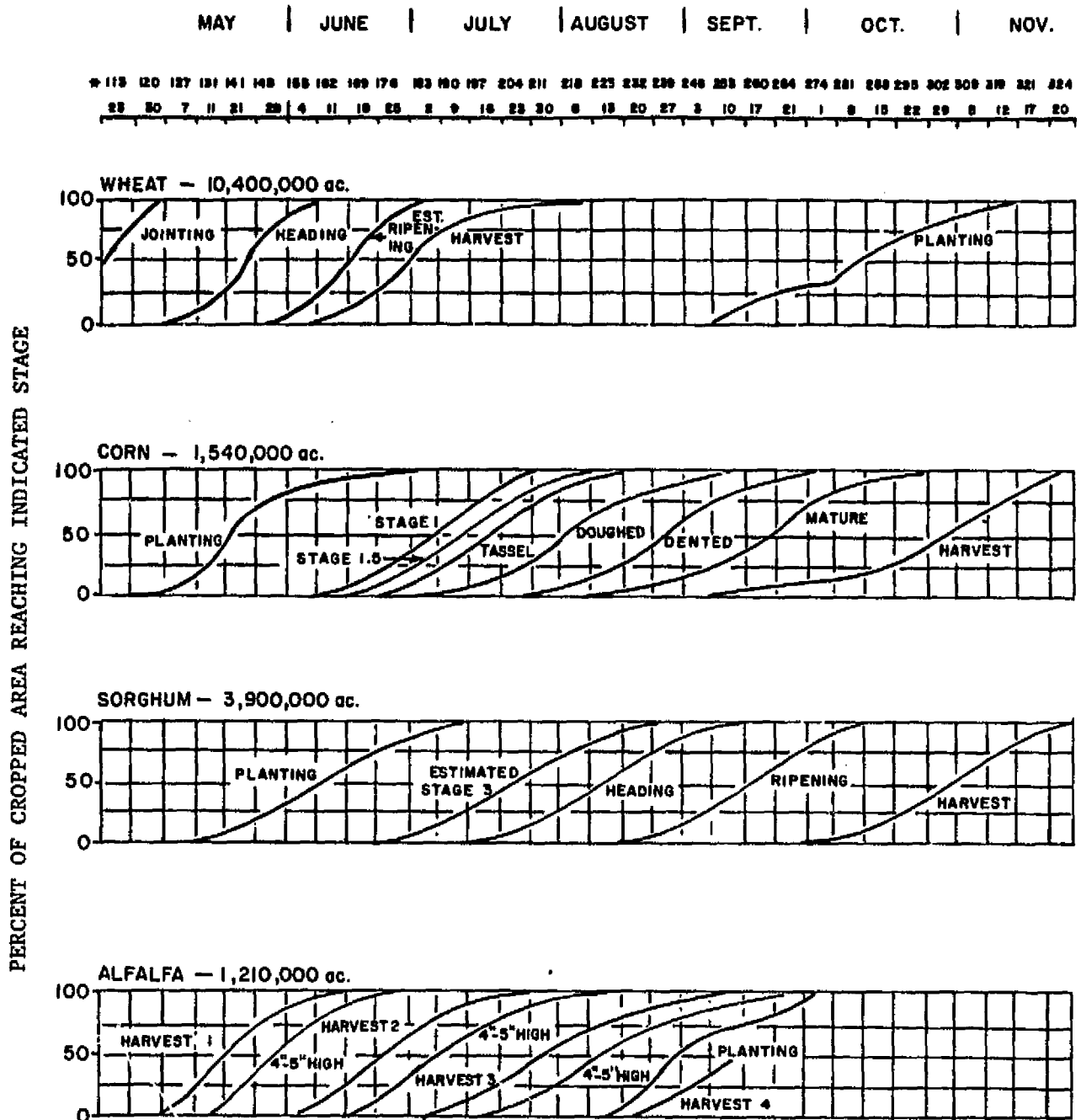


* JULIAN DATES

Source:  U.S. Weekly Crop and Weather Bulletin

The experimental findings suggest that the driver of discrimination is not spectral shape, rather the relative difference of amplitude of the spectral mean values, which occurs when the crop species are observed at different growth stages.

This does not connote the absence of all information in the quantitative differences between diverse spectral bands. It simply says that the information contained within the inter-band variations is relatively less important than that inherent in the differences between the intensities of the mean reflectances. Stated in other words, if the spectral reflectance shapes of various crop or of a single crop at different growth stages, are normalized to a common average reflectance, the ability to discriminate between them is limited. Interband differences probably do arise, especially if the entire region of interest is imaged simultaneously, from momentary differences of leaf orientations, different fraction of underlying ground covered by the leaf canopy, differences in the reflecting properties of the underlying ground, and similar phenomena depending on geometry or upon electromagnetic wave impedances and transmissivities presented by the pattern of reflecting objects.

Nevertheless, the experimental data appears to indicate that attempts at exploiting differentials of spectral shapes in and by themselves is akin to attempting to distinguish targets from clutter using amplitude patterns alone. The resulting accuracies are limited.

The presence of the statistical fluctuations, typical of clutter phenomena, further reduces the discrimination capability. The phenomenology appears formally equivalent to the glint effect observed in radar.

The situation suggests, as a working hypothesis, a model equivalent to the bistatic radar mechanism. The sun provides the illuminator, the illuminated crop represents a "scintillating" target, and the sensor is the receiver. In this model, the discrimination error is analogous to a false alarm rate.

The hypothesis leads to the development of a generalized theory of discrimination, presented in Chapter 5 following.

## 5.0 THEORY OF DISCRIMINATION FROM RADIOMETRIC DATA

### 5.1 Simplified Theory

The objective of agricultural acreage mensuration is to determine the proportion of the crop or crops sought relative to the total area under analysis, independently of where each crop species is located geographically. A first approach to the theory can be made by invoking the following assumptions:

- Two species only; normal distribution of reflectances

- One radiometric band only

- Area under study sufficiently large so that mensuration errors induced by the finite size of the resolution element (pixel) are negligible

It is shown in Appendix E that, under these assumptions, the error attributable to discrimination in estimating the proportion of the crop sought is given by:

$$e_d\% = \frac{100}{p\sqrt{n}} \frac{k}{(\rho_A - \rho_B)} \sqrt{p\,\sigma_A^2 + (1-p)\,\sigma_B^2} \qquad (5-1)$$

where:

   $p$ = proportion of crop A to total area

   $e_d\%$ = percent error in p attributable to discrimination

   $k$ = statistical confidence multiplier

   $n$ = total number of pixels in scene (species A pixels + species B pixels.

$\rho_A$, $\rho_B$ = means of the relative reflectances of species A, B respectively.

$\sigma_A$, $\sigma_B$ = standard deviations of the fluctuations of $\rho_A$, $\rho_B$ respectively.

Expression (5-1) can also be written:

$$e_d \% = \frac{100\, k\, h_A}{p\, \sqrt{n}\, (1 - \frac{\rho_B}{\rho_A})} \sqrt{p - (\frac{\rho_B}{\rho_A})^2\, (1-p)\, (\frac{h_B}{h_A})^2} \qquad (5\text{-}2)$$

where:

$$h_A = \frac{\sigma_A}{\rho_A} = \text{coefficient of variation of } \rho_A$$

$$h_B = \frac{\sigma_B}{\rho_B} = \text{coefficient of variation of } \rho_B$$

all other symbols are as defined previously

$$\frac{\rho_B}{\rho_A} = \text{is the } contrast\ ratio \text{ between the reflectances of the two crops}$$

## 5.2 Significance

There is currently no definitive experimental indication that the coefficients of variation differ systematically and reliably among crops. Until more comprehensive data become available, it is assumed that $h_A = h_B = h$.

Expression (5-2) can thus be modified by posing $h_A = h_B = h$. This assumption converts (5-2) into:

$$e_d\% = \frac{100\ k\ h}{p\ \sqrt{n}\ (1 - \frac{\rho_B}{\rho_A})} \sqrt{p + (1-p)\ (\frac{\rho_B}{\rho_A})^2} \qquad (5\text{-}3)$$

Explicitating n from (5-3) yields:

$$n = \frac{10^4\ h^2\ k^2}{p^2\ (e_d\%)^2\ (1 - \frac{\rho_B}{\rho_A})^2}\ [p + (1-p)\ (\frac{\rho_B}{\rho_A})^2\ ] \qquad (5\text{-}4)$$

Let us survey briefly the implications of this expression.

p is a characteristic of the target crop, in fact it is the quantity to be measured. p does vary from year to year: in most practical cases, the variation is relatively small. More significantly, p lies beyond the observer's control. He can however render the computation of p more facile *if he is privy to certain prior information.*

$e_d\%$, the allowable one-sigma error, is established by the user requirements. It is also beyond the observer's control.

$\rho_A$, $\rho_B$ and h are determined by the characteristics of the crops being inventoried. Control over $\rho_A$, $\rho_B$ is possible by selecting the timing of the observation.

n represents the number of pixels required to be observed. It is analogous to the "number of looks" in radar theory. It is controllable by varying the extent of the surface area being observed.

n is a driver of the characteristics and cost-effectiveness of the Data System, because:

1. The Data System's load is directly proportional to the number of pixels to be observed.

2. The costs of the data relay, preprocessing and information extraction portions of the Data System grow in proportion to the data load.

A feel for the relationships between parameters is available from Figure 5-1. It plots the total area required to be observed and measured versus the contrast ratio $\frac{\rho_B}{\rho_A}$, for the typical conditions: error desired 2%; confidence one sigma (k=1); p≈0.15 (a typical realistic value for densely cropped areas); and a linear pixel dimension of 30 meters (proposed for LANDSAT D). As shown in Chapter 4, the values of h used as parameters, e.g. 0.1 and 0.3, are typical of minimum and maximum values competing to realistic agricultural scenes.

Figure 5-1 can be converted to the characteristics of LANDSAT A, B, C, by increasing the ordinates by the square of the ratio of the respective resolutions, i.e. approximately fivefold.

A feel for typical practical values of the contrast ratio $\frac{\rho_B}{\rho_A}$ can be obtained from Figure 5-2, derived from the experimental data summarized in Chapter 4. It shows that under conservative system regimes, taking into account compression effects induced by the atmosphere, one should not assume design goals for $\frac{\rho_B}{\rho_A}$ significantly lower than 0.8; and that 0.95 represents an approximate upper bound, beyond which the area which must be measured generally transcends practically attainable limits.

# FIGURE 5-1

## AREA WHICH MUST BE OBSERVED AS A FUNCTION OF SCENE PARAMETERS

DISCRIMINATION ERROR, $e_d$ = 2%

PROPORTION  p = 0.15

RESOLUTION  r = 30 METERS

AVERAGE U.S. COUNTY

TYPICAL
OPERATING
REGION

$h = 0.3$

$h = 0.1$

h = COEFFICIENT OF VARIATION OF REFLECTANCE

AREA REQUIRED TO BE OBSERVED, $KM^2$
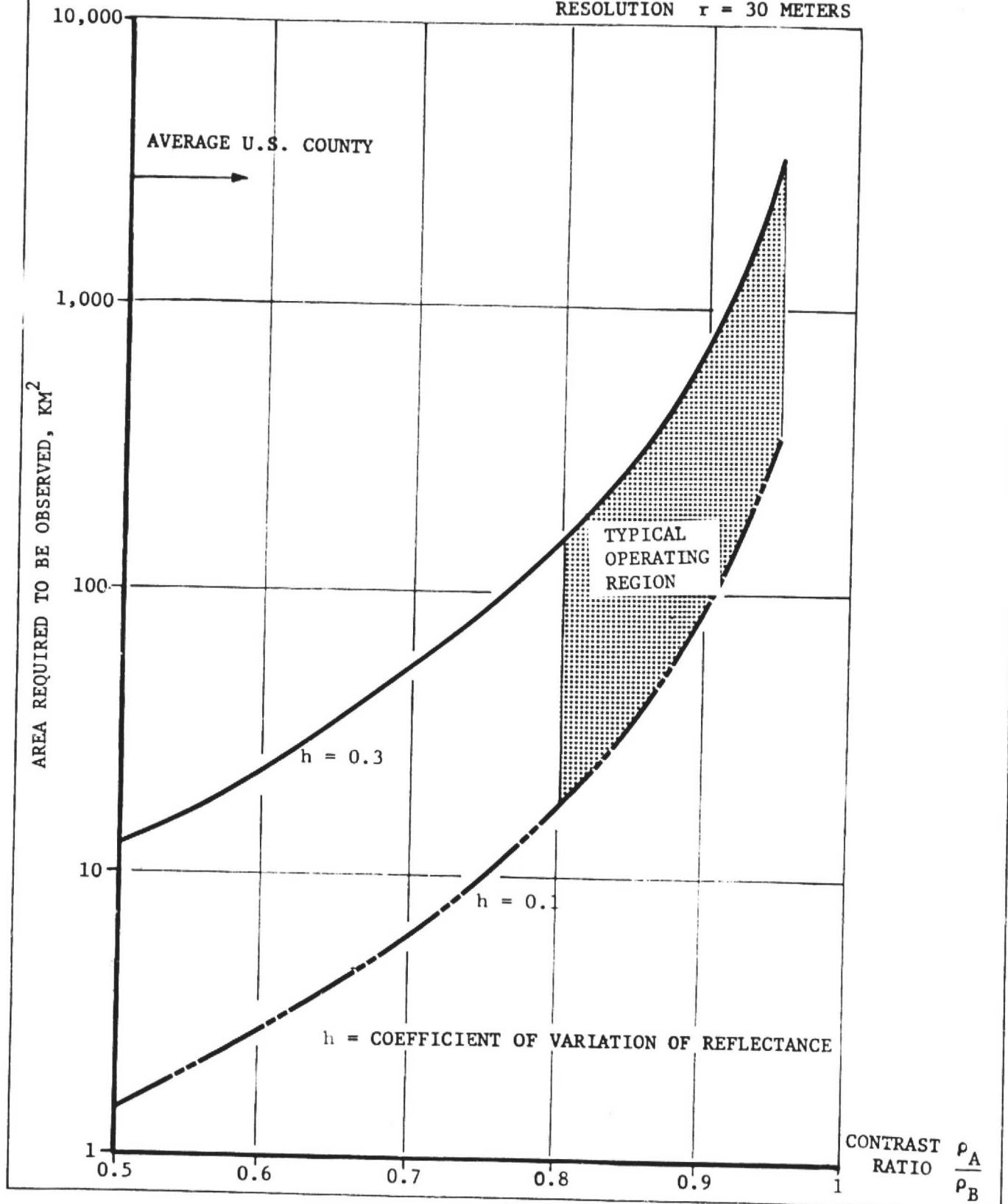
CONTRAST RATIO $\frac{\rho_A}{\rho_B}$

FIGURE 5-2

TEMPORAL BEHAVIOR OF CONTRAST RATIO: HAND COUNTY, N.D.

Contrast Ratio $\frac{\rho_B}{\rho_A}$
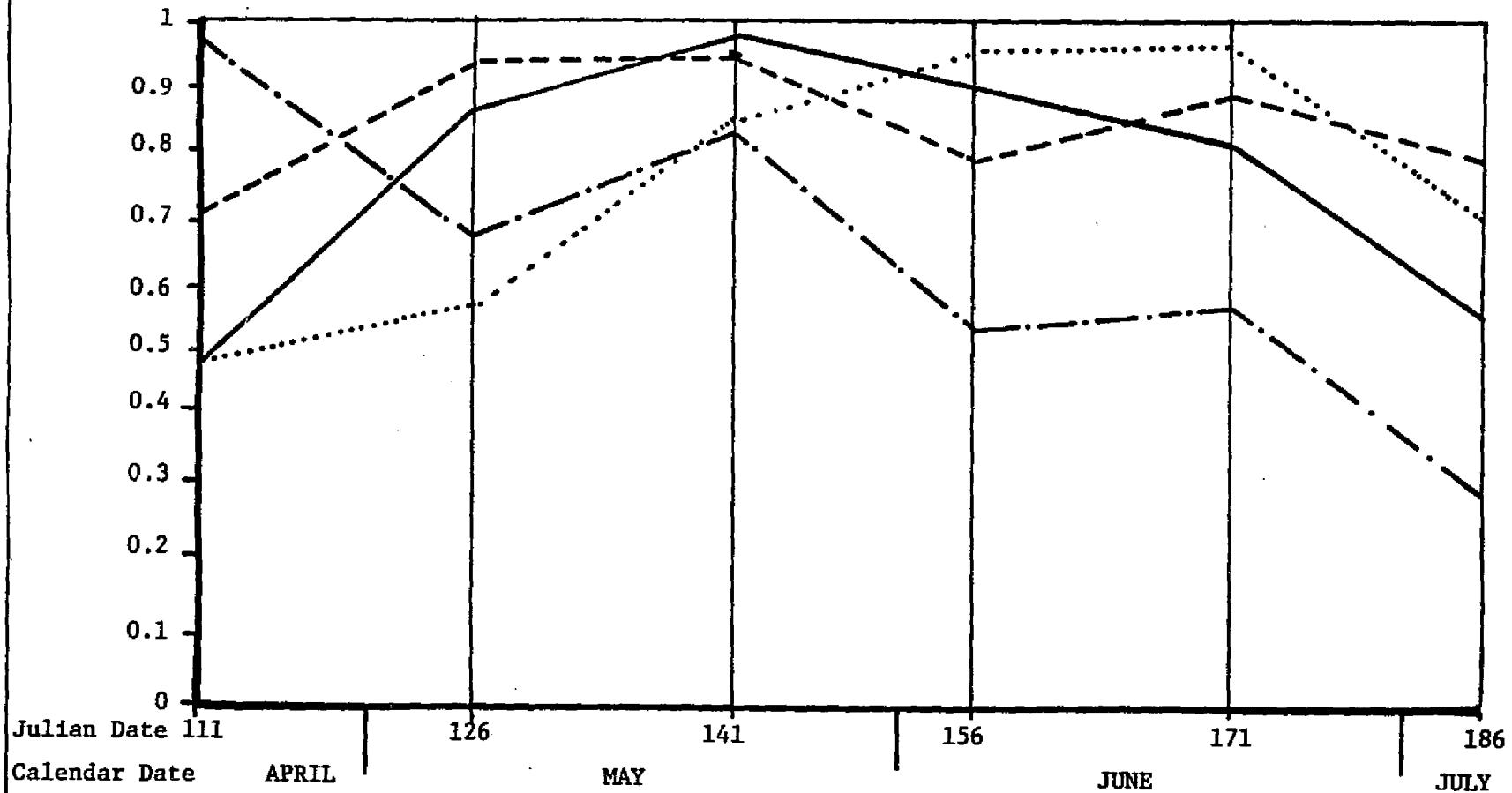
——Barley vs. Wheat    ······· Oats vs. Wheat    —·—· Alfalfa vs. Wheat    ———— Pasture vs. Wheat

Note: All wheat is spring wheat

## 5.3 Selection of the Best Band

The simplified theory deals with a single band. Given a choice among various bands, it comes naturally to ask which band is indeed the best and how to select it.

The most efficient system is obviously the one which provides the best performance at minimum cost. From the aforegoing, this is tantamount to require the simultaneous minimization of the error and of the number of pixels which must be sampled. From expression (5-2), it is clear that the product $e_d\% \sqrt{n}$ can be taken as representing a figure of merit of the Data System. Thus, the "best" single band is the one which, for specified conditions of observation, minimizes the figure of merit $e_d\% \sqrt{n}$. This is equivalent, from (5-2), to minimizing the expression:

$$e_d\% \sqrt{n} = K \frac{h_A}{1 - \dfrac{\rho_B}{\rho_A}} \sqrt{p + (1-p) \left(\frac{\rho_B}{\rho_A}\right)^2 \left(\frac{h_B}{h_A}\right)^2} \qquad (5-5)$$

where the constant K subsumes all parameters not affected by the choice of band. Under the previously introduced assumptions $h_A = h_B = h$, (5-5) becomes:

$$e_d\% \sqrt{n} = K \frac{h}{1 - \dfrac{\rho_B}{\rho_A}} \sqrt{p + (1-p) \left(\frac{\rho_B}{\rho_A}\right)^2} \qquad (5-6)$$

In other words, the "best" band consistently exhibits combinations of $\rho_A$, $\rho_B$, $h_A$, $h_B$ which minimize expression (5-5) — or (5-6) if one does not have available the separate values of $h_A$, $h_B$.

Tests were run to determine the best band or bands from the available experimental data. To this effect, pairs of crops were chosen whose ground cover reaches substantial values, in excess of 50% during the season. The reflectance values of the fine-grained spectra were integrated within the wavelength intervals of the first five TM bands.

From (5-6), the number of pixels required to be observed at each of the available temporal dates were calculated. Typical experimental results are shown in Table 5-1 for $h = 0.3$, error $e_d = 2\%$, $p = 0.15$, crops alfalfa versus barley, Hand County data. It is clear that the bands which require the lowest number of pixels, i.e. which minimize the "cost function," are the best.

TABLE 5-1

NUMBER OF PIXELS (COST FUNCTION) REQUIRED TO BE OBSERVED WITHIN THE FIRST FIVE TM BANDS

| TM BAND | WAVELENGTH, μm | DATES, (1977) 4/21 | 5/10 | 6/1 | 6/16 | DATES, (1977) 4/21 | 5/10 | 6/1 | 6/16 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.45 - 0.52 | - | $2.9 \times 10^5$ | $2.9 \times 10^6$ | $2.3 \times 10^4$ | - | 34 | 345 | 2.7 |
| 2 | 0.52 - 0.60 | $8.3 \times 10^5$ | $1.7 \times 10^6$ | $4.6 \times 10^7$ | $4.9 \times 10^4$ | 99 | 202 | 5,480 | 5.8 |
| 3 | 0.63 - 0.69 | $1.6 \times 10^6$ | $1.9 \times 10^6$ | $1.0 \times 10^6$ | $6 \times 10^4$ | 190 | 226 | 119 | 7.1 |
| 4 | 0.76 - 0.90 | $3.2 \times 10^4$ | $2.6 \times 10^4$ | $3.2 \times 10^5$ | $8.4 \times 10^3$ | 3.8 | 3.1 | 38 | 1 |
| 5 | 1.55 - 1.75 | $4.5 \times 10^7$ | $2.7 \times 10^5$ | $4.0 \times 10^5$ | $2.5 \times 10^5$ | 5,360 | 32 | 48 | 30 |

NUMBER OF PIXELS REQUIRED FOR DISCRIMINATION — COST FUNCTION NORMALIZED TO LOWEST VALUE

The example of Table 5-1 shows that TM Band 4 is the best *for the specific case selected.*

In general, the best band is case-dependent.  Various tests run for diverse crops at 50% or better ground cover resulted in the orientative rankings of the discriminating power of the TM bands depicted in Table 5-2.

TABLE 5-2

RANKING OF TM BANDS FOR DISCRIMINABILITY BETWEEN STAPLE CROPS

| TM BAND | WAVELENGTH, MICRONS | APPROXIMATE RELATIVE FIGURE OF MERIT |
|---------|---------------------|--------------------------------------|
| 1 | 0.45 - 0.52 | 0.1 |
| 2 | 0.52 - 0.60 | 0.1 |
| 3 | 0.63 - 0.69 | 0.15 |
| 4 | 0.76 - 0.90 | 1 |
| 5 | 1.55 - 1.75 | 0.3 |

## 5.4  Effect of Multiple Bands

Rigorous extension of the theory to multiple bands was not attempted in this effort:  because the objective was to determine the drivers of Data System performance rather  than to develop discrimination theory. Some general considerations on the effect of multiple bands can however be made at this time:

1.  The addition of other bands to the best band will generally improve the discriminant power of the system; provided, that the signal-to-fluctuation ratio this added to the best band is not so low as to

yield a decreased overall signal-to-fluctuation ratio for the
multiple bands.

2. The addition of bands is beneficial - subject to the caveat ex-
pressed above - even if their spectral amplitudes are completely
correlated (cf. Table D-2, Appendix D).  This is because adding
bands provides additional samples.  Thus, if all bands possess
the same discriminating power, the upper bound of the improvement
is given by the sampling theorem, i.e. it will equal $\sqrt{b}$, where
b is the number of bands combined.  In practice, as indicated by
Table 5-2, the benefit will be lower than this upper bound, if
the best band has been selected initially.

In conclusion, all other conditions remaining invariant, multispectral
discrimination can be expected to lower the number of pixels and corres-
ponding total area required to be observed (cf. Figure 5-1), but not by
a strikingly large amount.

# 6.0 CONSEQUENCES OF THE THEORY ON THE CHARACTERISTICS OF THE DATA SYSTEM

## 6.1 Impact on the Frequency of Observation

The developments of previous Chapter 5 lead to the following conclusions regarding accurate discrimination:

1. It is not sufficient to image agricultural scenes "when all crops look good." They must be imaged when the statistical "distances" between their probability distributions reach sufficiently large values. The requires selecting moments of observation for which $\frac{\rho_B}{\rho_A}$ and $h_A$, $h_B$ are, simultaneously, sufficiently small.

2. The above conditions do not occur for all crops simultaneously: rather, they occur generally pairwise. The net result is that quality discrimination must be achieved by analyzing several optimally positioned temporal "cuts."

The typical situation is depicted in Figure 6-1. The time intervals during which the "statistical distances" are widest bear the generalized designation of "temporal windows", or simply "windows."

The duration of a window is not a univocal quantity. It is a function of:

1. The system's specified figure of merit, e% $\sqrt{n}$. The more stringent the figure of merit --- i.e., the higher the desired accuracy and simultaneously the lower the number of pixels one wishes to process --- the smaller must be the value of function (5-5). Consequently the shorter will be the admissible window. In extreme

# FIGURE 6-1

## CONCEPT OF STATISTICAL WINDOW

DATA FROM HAND CO., N.D.

SPECTRAL BAND:  0.76-0.90
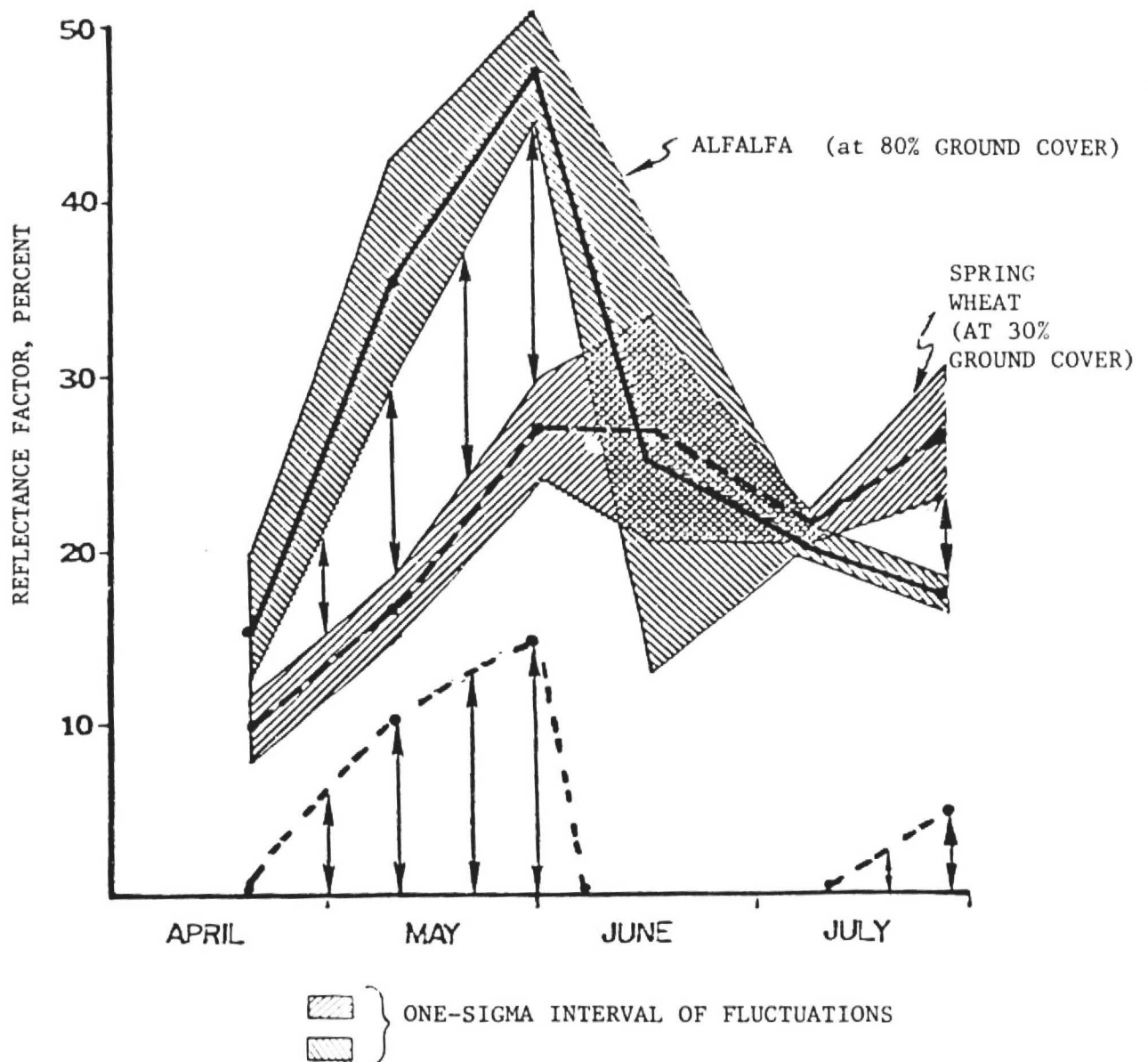
cases, i.e. if the specification of the figure of merit is made
too stringent, the window may vanish altogether.

2.  The system's sensitivity threshold $(\frac{\rho_B}{\rho_A})_{th}$. The previous deriva-
    tions disregard exogenous fluctuations acting within the scene.
    In practice, such fluctuations exist. Thus it is necessary in
    real situations to allow a "margin of safety" on the threshold
    value of $\frac{\rho_B}{\rho_A}$. In other words, the actual $\frac{\rho_B}{\rho_A}$ must be chosen some-
    what more stringently, or lower, than its theoretically accept-
    able value.

    The principal exogenous fluctuations stem from three sources:

    o  The granularity of the sensor, caused by quantization
       errors, calibration imperfections, instrumental errors.

    o  The variability of the atmospheric absorption across the
       scene.

    o  The variability of the path radiance across the scene.

    As regards the sensor, a reasonable estimate, derived from MSS
    performance, is to assume that the theoretical maximum number
    of levels is reduced by a factor of two. Thus, for example the
    Thematic Mapper, which possesses a theoretical maximum of 256
    levels, probably contains only approximately 100 univocal levels
    at a false alarm ratio of order 1%.

    Because experimental results show that the reflectance fluctua-
    tions of crop targets, also known as "scene noise," are consi-
    derably larger than this value, sensor fluctuations are negli-
    gible to the effects of performing agricultural inventories by
    radiometric remote sensing from OSTA space platforms.

    Atmospheric transmissivity fluctuations affect $(\frac{\rho_B}{\rho_A})_{th}$ are those
    which occur between diverse points of the scene. Variations
    which affect the entire scene uniformly affect both $\rho_A$ and $\rho_B$
    by the same ratio and thus impact only slightly the ratio
    $(\frac{\rho_B}{\rho_A})_{th}$.

Information on the intra-scene fluctuations is limited. Investigations by ECOsystems, using data from the US Turbidity Network indicate that fluctuations between separate points of the scene can reach peak excursions as high as 20%, rms values of as much as 10%. Figure 6-2. The contribution of these fluctuations to the uncertainty in $(\frac{\rho_B}{\rho_A})_{th}$ depends upon the distribution of the species A, B across the scene. If A and B are adjacent to each other everywhere (locally compact), the fluctuations will have no effect on $(\frac{\rho_B}{\rho_A})_{th}$. If however the locations of A and B are disaggregated (for example, as a limiting case, all the A's are in one place, all the B's in another, distant from the first), then $\rho_A$ and $\rho_B$ will suffer different random variations. Pending the availability of supplementary information on atmospheric fluctuations, it is not unreasonable to assume that $\rho_A$, $\rho_B$ are both affected by uncorrelated random fluctuations of 3%: their combined rms fluctuation thus would be of order 4%, one sigma.

The third factor, variability of the path radiance, is even less known. It is not included in this calculation.

3. The amount of prior information available. If the phenologic progress for the year under study is poorly known, the average phenologic progressions are computed from historical statistics. Figure 6-3 illustrates the procedure for determining the temporal fluctuations of the phenologic stages from historical data, in this case from [10]. It is noted that the yearly progressions of mean reflectances are shown as reaching the same peak values. This is not necessarily true in nature: it is an assumption dictated by the scarcity of experimental data. In practice, it can be expected that the peak values will vary from year to year, in response to differing situations of ground cover, plant stress, and similar. The computation of the temporal excursions is however affected but little by fluctuations of reflectance levels.

FIGURE 6-2

MONTHLY MEAN AND ONE-SIGMA DISPERSION OF AVERAGE DAILY

TRANSMITTANCE DIFFERENTIAL

TAKEN AT 5-DAY INTERVALS, BETWEEN RALEIGH AIRPORT AND
RESEARCH TRIANGLE PARK, NORTH CAROLINA



Stations located 8 km apart.

All observations taken on
cloud-free days or when a
clear path to the sun is
available.

σ (one-sided)

Average

# FIGURE 6-3

## PROCEDURE FOR CONSTRUCTING PHENOLOGICAL STATISTICS

RELATIVE
REFLECTANCE

TM BAND 4

KANSAS

1973

1972

1971

+ SEVEN ADDITIONAL PAST YEARS

he = HEADING
ts = TASSELING
pod = PODDING
● = EMERGING

WHEAT ————
CORN — — —
SOYBEANS —·—

10-YEAR AVERAGE

-σ →        ← +σ

APR    MAY    JUN    JUL    AUG    SEP    OCT    NOV

4. The presence of cloud cover. The random occurrence of clouds during the window intervals can be calculated or simulated with existing cloud models. This is further discussed in Chapter 7.

Figure 6-4 depicts typical windows. Note the effect upon window length of varying the sensitivity threshold of the system. The windows discussed previously are for cloudless conditions. Figure 6-5 depicts the required frequency of observation for selected regions of the world when cloud cover is taken into account for various levels of threshold sensitivity. Chapter 7 illustrates how the windows portrayed in Figure 6-4 for selected regions were expanded worldwide by simulation. The results were analogous to those portrayed in Figure 6-5: recurrence times of order 5 days are needed to achieve coverages commensurate with 96% to 98% completeness of crop acreage mensuration.

Chapter 8 discusses ways and means of achieving such low recurrence rates with minimal cost impact upon the system.

## 6.2 Impact on the Extent of the Area which must be Observed

The preceding theory shows that it is not necessary to observe, sample and process the entire surface of the region under study. If the times of observation are selected correctly, sampling of considerably smaller portions is sufficient. Let us examine this in more detail.

### 6.2.1 The Effect of Bounding and the Role of Auxiliary Information

The sensitivity of the area to be measured to its crop proportion is depicted in Figure 6-6, which also presents for comparison the propor'ions

FIGURE 6-4

STATISTICAL (ONE SIGMA) TEMPORAL WINDOW ENVELOPES OF REFLECTANCE DIFFERENCE

USSR, UKRAINE
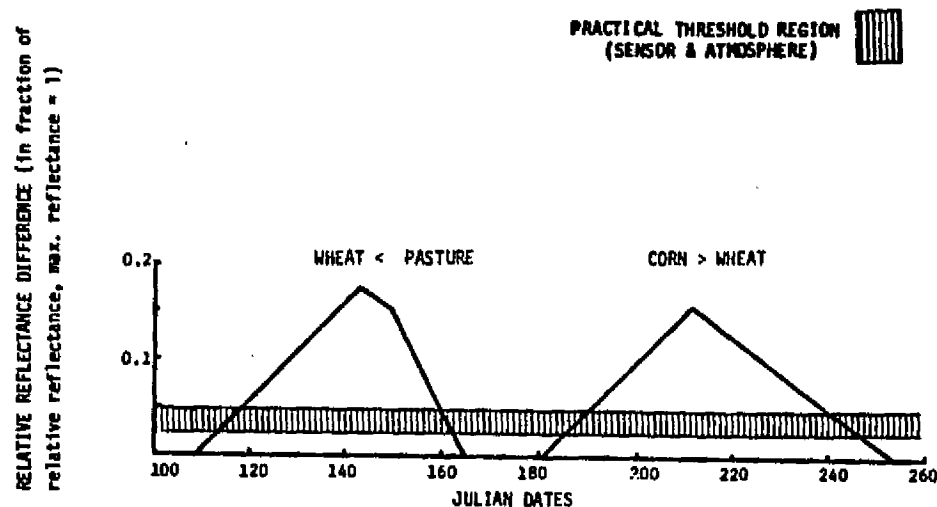


FIGURE 6-5

REQUIRED RECURRENCE OF OBSERVATION TO ACHIEVE 99% PROBABILITY OF VIEWING FOR SPECIFIED
CLOUD COVER AS A FUNCTION OF SENSOR THRESHOLD SENSITIVITY $\Delta\rho$
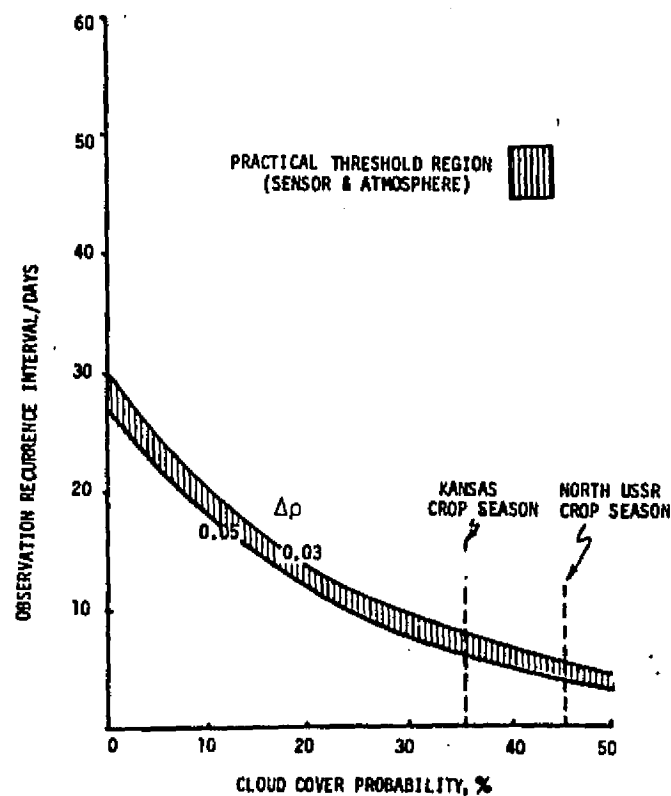
FIGURE 6-6

DEPENDENCE OF AREA TO BE OBSERVED ON THE PROPORTION

OF THE CROP TO BE INVENTORIED

ERROR DESIRED = 2%, 1$\sigma$

PIXEL RESOLUTION = 30 meters
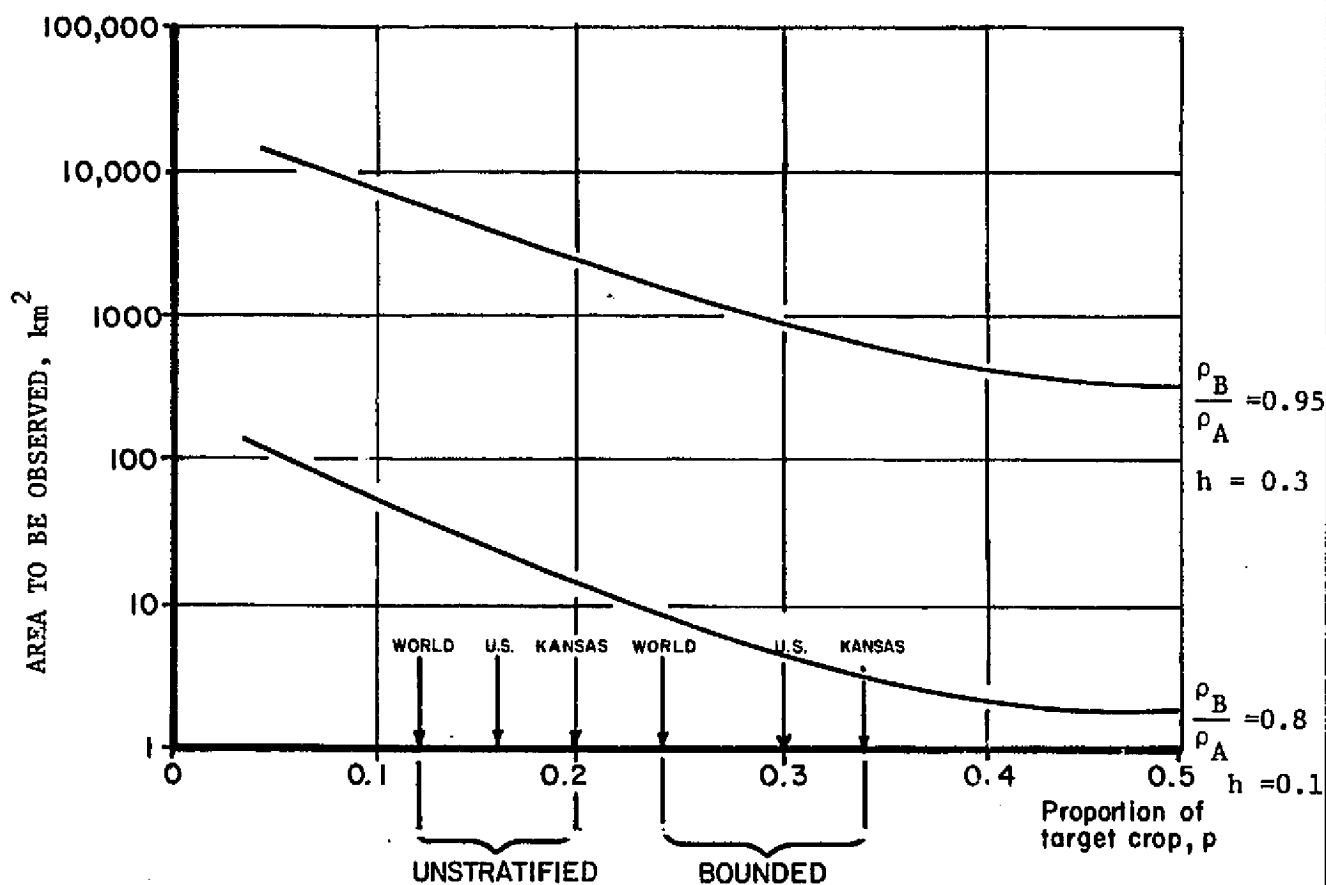
competing to selected areas of the world. It can be seen that the average
proportions of crops are relatively small, calling for the sampling of
rather extensive areas: but that the area to be observed decreases dras-
tically as the crop proportion increases.

This leads to the concept of "bounding," i.e. eliminating from the
observed scene as many as possible areas known not to contain the crop
or crops sought. The concept of bounding is obvious for such surface fea-
tures as lakes, deserts, dense urban settlements. For other than these
obvious features, proper bounding requires the availability of prior in-
formation, i.e. what areas under observations do or do not contain the
crops sought. After bounding is accomplished to the best extent possible
with the information available, stratification into diverse levels of crop
proportions is effective in further reducing the area to be observed.
Same as with bounding, proper stratification also requires prior informa-
tion.

### 6.2.2 Sampling

If the moments of observation are correctly chosen, the area required
to be observed will be significantly less than the net scene area still
available, even after bounding and stratification. In other words, it
will not be necessary to observe all of the residual area: observation
of a smaller fraction generally suffices to yield the desired accuracies.
(5-4) allows computation of the area required to be observed: however,
it does not provide information as to what geographical locales within
the scene should be observed.

In the absence of other prior information, the distribution of the locales to be observed is determined by statistical sampling. The standard procedure is to choose the number of samples, or segments, from the well-known formulation:

$$e_s = \frac{\tilde{\sigma}}{\tilde{\mu} \sqrt{N}} \tag{6-1}$$

where:

$e_s$ = sampling error desired in the measurement of p

$N$ = number of sample segments

$\tilde{\sigma}$ = fluctuation of proportion within the sampled area

$\tilde{\mu}$ = mean value of the proportion measured from the N samples.

In practice, $\sigma$, or, equivalently, the coefficients of variation $\frac{\sigma}{\mu}$ of agricultural regions are not published nor available. What is available is the historical sequence of the region's proportions of target crops p. If the successive p's do not fluctuate too widely, one can assume their temporal average as representative of the region: otherwise, in absence of other information, it is prudent to assume the lowest value of p experienced over the last several years.

This historical value of p can be utilized in place of the unavailable $\frac{\sigma}{\mu}$ by assuming that the population of samples to be chosen obeys a Bernoullian distribution, wherein each sample, or segment, is counted as zero if it contains a proportion of the target crop less than p: and as 1, if more. With this assumption, the standard deviation $\tilde{\sigma}$ is given by:

$$\tilde{\sigma} = \sqrt{N \, p \, (1-p)} \qquad (6-2)$$

The mean $\tilde{p}$ of the sample is $\dfrac{\sum\limits_{1}^{1,N} p_i}{N}$, which, for N sufficiently large, tends to p. The mean proportion containing the species sought is then Np, and expression (6-1) becomes:

$$e_s = \frac{\sqrt{Np \, (1-p)}}{Np} = \sqrt{\frac{1-p}{p}} \ \frac{1}{\sqrt{N}} \qquad (6-3)$$

Expressing $e_s$ as a percent error, and introducing the confidence multiplier k, (6-3) becomes:

$$e_s = 100 \ k \ \sqrt{\frac{1-p}{Np}} \qquad (6-4)$$

This expression is valid if Npk >> 1, i.e. to the extent to which the Bernoullian approximation is close to the Gaussian distribution, and if N << total number of sample segments extractable from the observed scene. In practice, the errors are negligible at the 1% as long as Npk $\geq$ 30, and if N/(Parent population N) $\leq$ 2%.

The formulation above applies to unstratified sampling. The effect of stratification is to increase the "effective" p by parceling the samples according to the p - density of the various strata within the scene.

Expression (6-4), or its originating form (6-1), do not in and by themselves answer the question of how large to make each individual segment. As the segment's surface is made progressively larger, it

eventually will encompass the entire area to be inventoried. The number of segments becomes unity, yet the error becomes zero: thus (6-4) breaks down in this extreme case. A quantitative formulation requires knowledge of the distribution of p within the region. In its absence, (6-4) still holds for conditions wherein the total area sampled is "sufficiently" small with respect to the total area being inventoried, and the number of segments is "sufficiently" large.

From (6-4), the number of sample segments can be calculated:

$$N = \frac{10^4 \, k^2 \, (1-p)}{(e_s\%)^2 \, p} \tag{6-5}$$

We have seen that the total number of pixels required for discrimination is given by expression (5-4). The question is, how many pixels to allocate to each segment, i.e. what is the optimum area of each segment.

If there are no other error sources, and if the errors induced by sampling and those induced by discrimination are uncorrelated, then the total error is given by:

$$e_t = \sqrt{e_s^2 + e_d^2} \tag{6-6}$$

where $e_t$ is the inventory error specified by the user, $e_s$ is the sampling error, and $e_d$ is the discrimination error.

Expression (6-6) indicates that there is a choice in the selection of N, number of sample segments, and $n_t$, total number of pixels to be

observed: as long as neither of the corresponding errors $e_s$, $e_d$ equals the total error $e_t$, and subject further to the obvious constraint $N < n$.

For example, suppose that one wishes to minimize $n_t$, which is proportional to the data system costs. To see what happens, let us re-write (6-6) using expressions (6-4) and (5-3):

$$e_t^2 = \frac{h^2}{p^2 \, n_t \, (1 - \frac{\rho_B}{\rho_A})^2} \, [p + (1-p) \, (\frac{\rho_B}{\rho_A})^2] + \frac{1-p}{pN} \qquad (6-7)$$

Using for convenience:

$$\frac{h^2}{p^2 \, (1 - \frac{\rho_B}{\rho_A})^2} \, [p + (1-p) \, (\frac{\rho_B}{\rho_A})] \quad = A$$

$$\frac{1 - p}{p} = B$$

and substituting in (6-7):

$$e_t^2 = \frac{A}{n_t} + \frac{B}{N} \, , \qquad (6-8)$$

whence:

$$n_t = \frac{A}{e_t^2 - \frac{B}{N}}$$

To find the minimum $n_t$, since $e_t$, A, B are constants, set the condition:

$$\frac{dn_t}{dN} = \frac{AB}{N^2 \left(e_t^2 - \frac{B}{N}\right)^2} = 0 \tag{6-9}$$

Condition (6-9) is satisfied when the denominator tends to infinity, i.e. when $N \to \infty$.

Since, by definition, $N$ cannot exceed $n_t$, the minimum value of $n_t$ is achieved when $n_t = N$. From (6-8), $e_t$ becomes:

$$e_t = \frac{1}{\sqrt{n_t}} \sqrt{A + B}$$

In this case, each pixel acts both as a discrimination element and as a sample segment. The total number of required pixels will be given by:

$$n = \frac{A + B}{e_t^2} = 10^4 \frac{A + B}{(e_t \%)^2} \tag{6-10}$$

and will be somehwat larger than depicted in Figure 5-1. In practice, this mode of operation may be difficult to achieve by orbital sensors of the not too distant future, for two reasons: 1) it would entail a very rapid scan of the sensor during overpass; 2) it would entail high precision of pointing, to insure correct location of the "pixel segments" on the earth's surface.

On the other extreme, if one wishes to minimize the number of sample segments $N$, so as to minimize the sensor's angular excursions, it can be observed that $N$'s lower limit is given by (6-5). However, choice of

this value of N entails n = ∞, because it is tantamount to specifying $e_d$ = 0 in (6-6) and hence in (5-4). In this extreme case, the area of each segment would become infinite. Since this is not possible, the upper limit for this case is the already-discussed case of a single segment encompassing the entire area to be inventoried. As N is increased from this lower limiting value, n decreases monotonically until it reaches the value given by (6-10).

Optimization of the relative weights to be placed on $e_s$ and $e_d$ requires the formulation of an additional relationship between N and n. Such a relationship could be written, for example, if onw knew the improvement of sampling accuracy achievable by increasing the area subtended by each segment. This type of relationship would be a function of the characteristics of the observed region.

In the absence of such knowledge, one can assume that $e_s$ and $e_d$ possess equal weights in the total error budget:

$$e_d = e_s = \frac{e_t}{\sqrt{2}}$$

With this hypothesis, the total number of pixels required, $n_t$, is, from (5-3):

$$n_t = \frac{2 \times 10^4 \, h^2 \, k^2}{p^2 \, (e_t\%)^2 \, (1 - \frac{\rho_B}{\rho_A})^2} \, [p + (1 - p) \, (\frac{\rho_B}{\rho_A})^2] \tag{6-11}$$

and the number of samples, N, is from (6-5):

$$N = \frac{2 \times 10^4 \, k^2 \, (1 - p)}{(e_t \%)^2 \, p} \tag{6-12}$$

Consequently, the number of pixels per sample segment, $n_s$, becomes:

$$n_s = \frac{n_t}{N} = \frac{h^2}{p \, (1 - \frac{\rho_B^2}{\rho_A})} \, [\frac{p}{1 - p} + (\frac{\rho_B}{\rho_A})^2] \tag{6-13}$$

Figure 6-7 shows $n_s$ as a function of $p$ for values of the parameters $e_t$, $h$, $\frac{\rho_B}{\rho_A}$ corresponding to practical situations. The strong effect of increasing $p$ on reducing the sample areas is clearly evident. The corresponding number of segments, N, is shown as a function of $p$ in Figure 6-8.

As already mentioned, the computations above hold for unstratified sampling. Stratification generally reduces the number of samples required. USDA, for example, estimates (personal communications from Mr. William Wigton of USDA/ESCS) that the number of samples required to inventory U.S. wheat at a 2% sampling error is of order 1,000 with optimum stratification. This number would increase if one wished simultaneously to inventory additional crops.

The allocation of stratified samples is a function of the homogeneity of the areal distribution of acreage. To assess the latter, a map of the region's agricultural acreage is required. Because this varies with time, sampling schemes periodically require the availability of full frames, at intervals commensurate with the dynamics of the observed region: generally of order several years.

FIGURE 6-7

NUMBER OF PIXELS AND AREA OF EACH SEGMENT AS A FUNCTION OF PROPORTION

ERROR DESIRED =
= 2%, 1σ

$\frac{\rho_B}{\rho_A}$ = 0.9, h = 0.3

$\frac{\rho_B}{\rho_A}$ = 0.8, h = 0.3

$\frac{\rho_B}{\rho_A}$ = 0.8, h = 0.1

PROPORTION, p

AREA OF SEGMENT, ha.
@ 70 METER RESOLUTION

AREA OF SEGMENT, ha.
@ 30 METER RESOLUTION

NUMBER OF PIXELS
IN EACH SEGMENT
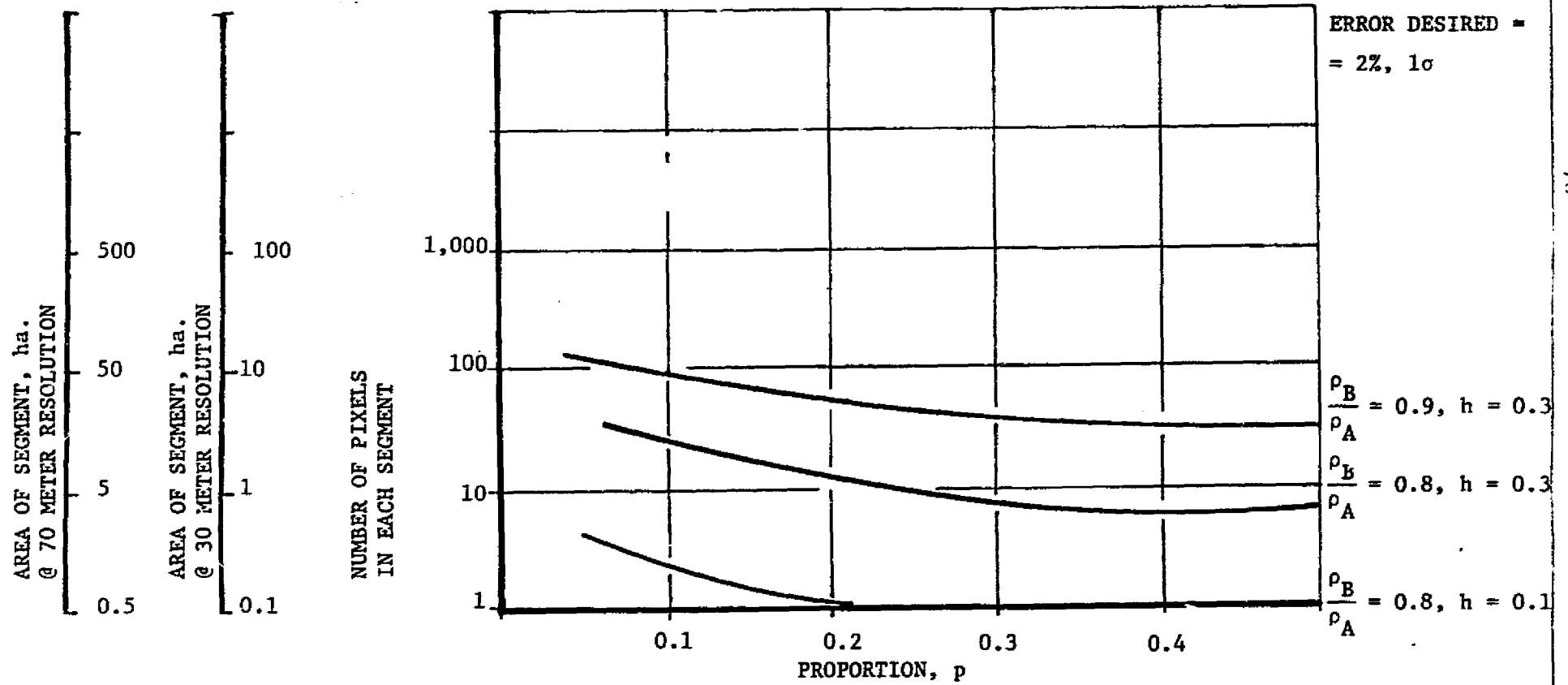
## FIGURE 6-8

### NUMBER OF SAMPLES REQUIRED TO ACHIEVE
### SPECIFIED ACCURACY AS A FUNCTION
### OF p
### UNSTRATIFIED SAMPLING
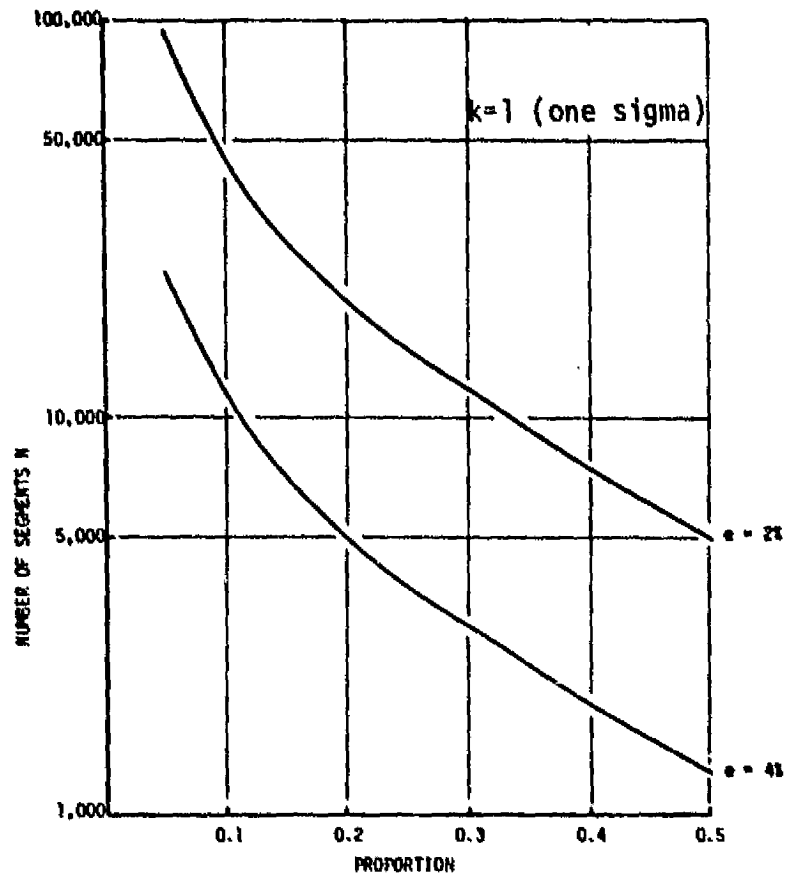


## FIGURE 6-9

### MENSURATION ERROR AS A FUNCTION OF TOTAL AREA SAMPLED

r = 30 meters, p = 0.15
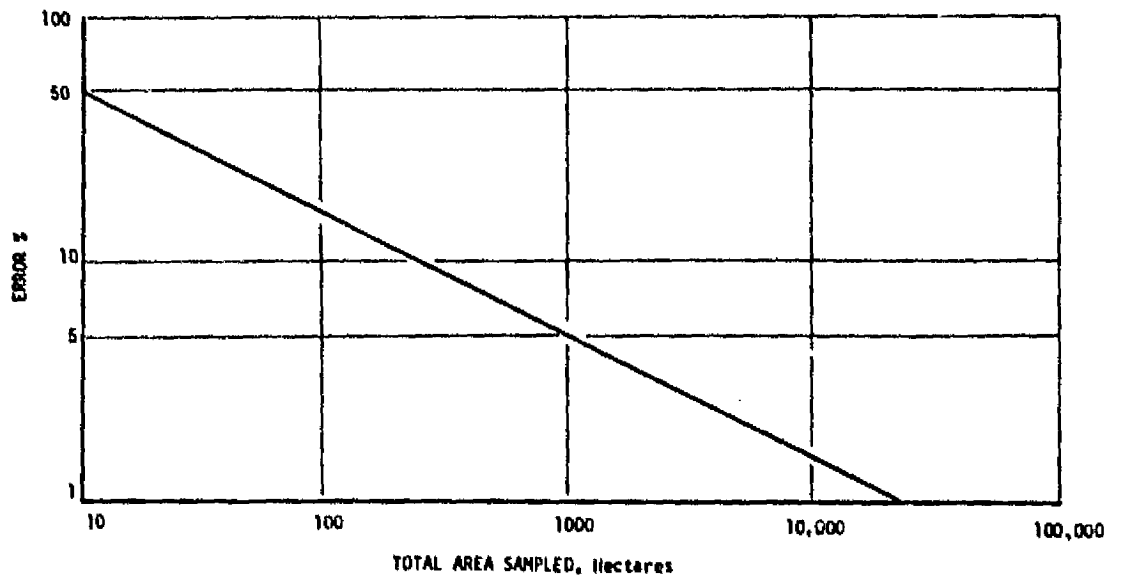
### 6.2.3 Mensuration

If the total area to be inventoried is relatively small, an additional source of error, well known in surveying practice, arises due to the finite precision of the measurement devices employed. In the case at hand, the precision is of order one pixel. The corresponding mensuration error is:

$$e_m\% \approx \frac{2\,\gamma\,\beta\,rk}{\sqrt{ApN}} \approx \frac{2\,\gamma\,\beta\,r\,k}{\sqrt{n\,p\,\delta}} \qquad (6-14)$$

Where:

$e_m\%$ = mensuration error, percent of crop area inventoried

$\gamma$ = form factor of sample (= 1 for square samples)

$\beta$ = mensuration figure of merit ($\approx$ 1 for photointerpreter)

$r$ = linear dimension of pixel, meters

$k$ = statistical confidence multiplier (= 1 for $1\sigma$)

$A$ = area of each segment, hectares

$p$ = proportion of crop area to total area

$N$ = number of sample segments

$n$ = total number of pixels in all sample segments

$\delta$ = area of each pixel, hectares (0.09 for 30 meter LANDSAT D pixels, $\approx$0.44 for LANDSAT C pixels).

The mensuration error, depicted in Figure 6-9 for practically encountered parameters, must be taken into account when it becomes a substantial fraction of the allowable total error $e_t$. This can happen if the total area to be sampled is small -- brought about for example, if p is very high and/or the contrast ratio $\frac{\rho_B}{\rho_A}$ is very low.

In this case the total error budget must be allocated between the three sources of error, i.e. sampling, discrimination, mensuration:

$$e_t^2 = e_s^2 + e_d^2 + e_m^2 \qquad (6-15)$$

In the absence of other prior knowledge, the error budgets are allocated equally. The simple manipulations result in expressions similar to (6-11), (6-12), (6-13), with the additional term due to mensuration error included. Figure 6-10 depicts the three sources of statistical error arising in proportion estimation, and the corresponding expressions for each and for the total error.

### 6.2.4  Advantages and Practical Constraints of Sampling

Sampling can significantly reduce the systems data load: specifically, in spacecraft-to-ground data relay, in preprocessing, and in the information extraction function.

The reduction theoretically achievable by sampling vis-a-vis wall-to-wall imaging is shown in Figure 6-11. In the near future, technological constraints limit its full realization. For example, the precision of the platform's stabilization system impacts the accuracy with which the sample segment can be located on the earth's surface. The typical stabilization accuracy of the Multi Mission Modular Spacecraft (MMS), planned standard support for LANDSAT-D, is ±0.25 arc min rms. From an orbital altitude of order 700 Km, the corresponding ground smear is of order ±140 meters, 99% of the times.

## FIGURE 6-10

RANDOM SYSTEM ERRORS IN ACREAGE MENSURATION

| SAMPLING ERROR $E_1$ | → | CLASSIFICATION ERROR $E_2$ | → | MENSURATION ERROR $E_3$ | → | $\underline{\text{TOTAL ERROR}}$ $\sqrt{E_1^2 + E_2^2 + E_3^2}$ |

o Caused by limited     • Caused by imperfect     o Caused by coarseness
  number of samples       discrimination         of pixel yardstick
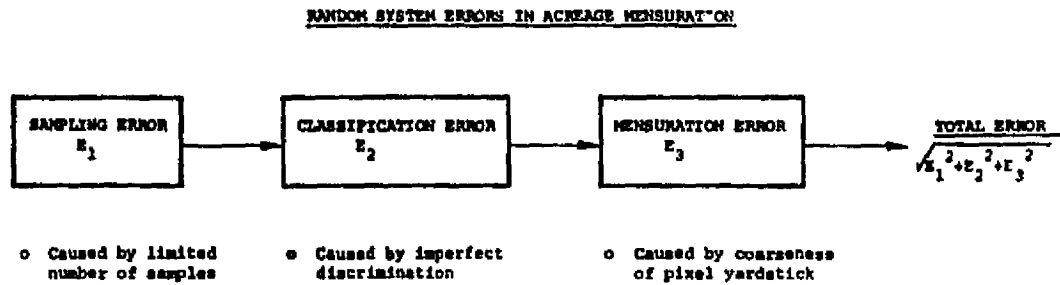
## FIGURE 6-11

### REDUCTION OF DATA LOAD THEORETICALLY ACHIEVABLE BY SAMPLING VERSUS WALL-TO-WALL COVERAGE



Pixel resolution = 30m

Wall-to-wall
coverage = 185 X 185 Km

Error = 2%

| $\dfrac{\rho_B}{\rho_A}$ | h |
|---|---|
| 0.9 | 0.3 |
| 0.8 | 0.3 |

DATA LOAD, PERCENT OF WALL-TO-WALL

Proportion, p

UNSTRATIFIED   BOUNDED
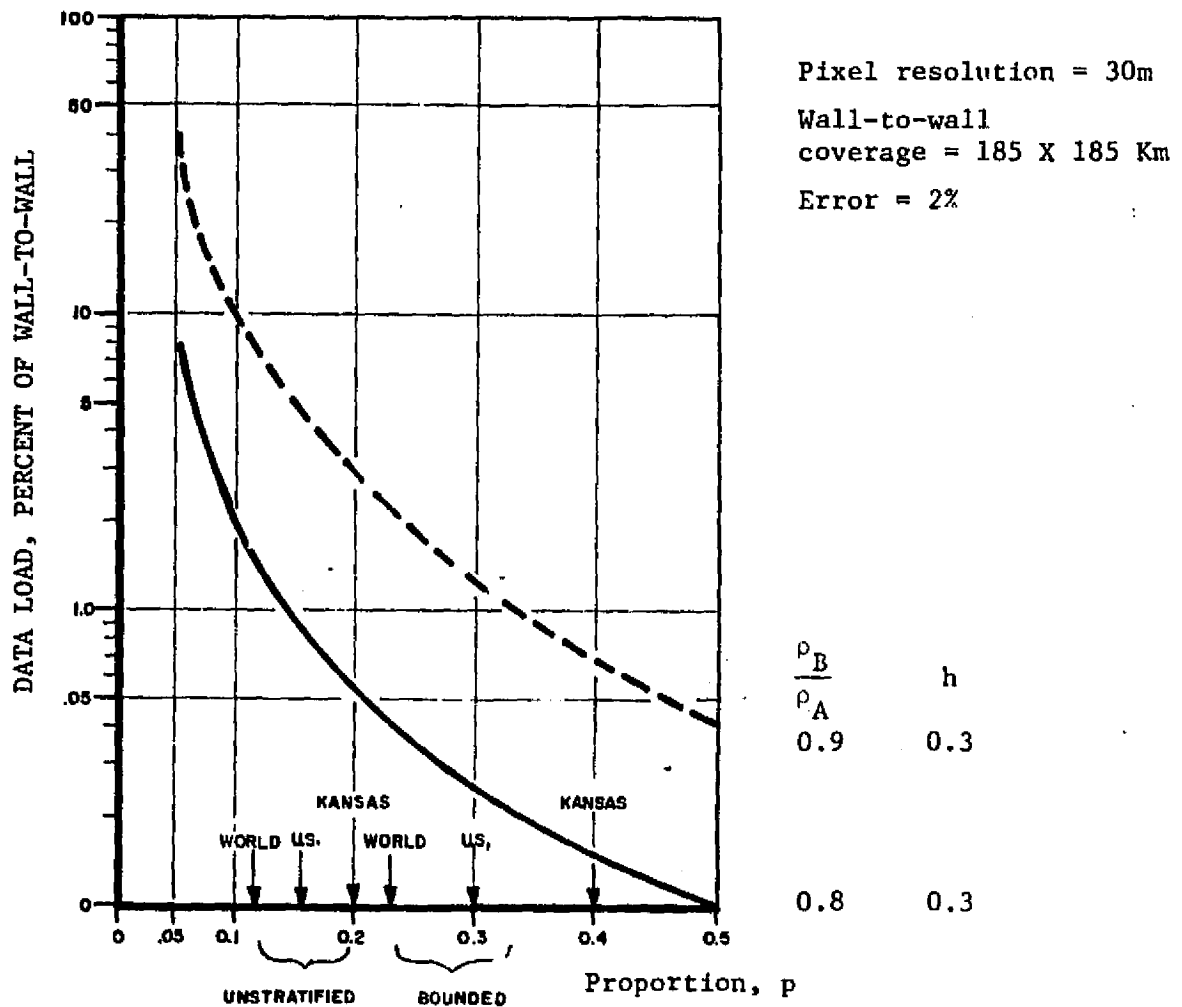
Sampling cannot be used to the total exclusion of all wall-to-wall data. The latter is still required periodically to update the stratification method.

## 6.3 Role of Auxiliary Information

The preceding developments illustrate how crop discrimination improves if prior available information is exploited in conjunction with radiometric data. In particular, the availability of prior information appears essential to the crop *identification* function.

Limited experimental results indicate the possible appearance of narrow band spectral features peculiar to particular species at some phenologic stages. The reliable presence of such "indicators" would allow identification without prior information. So far, the only such effect known to the authors to have been observed is the "red shift" of wheat, at the pre-flowering stage, found by S. Ungar of GISS [15].

The value of prior information to the discrimination process is greater, the less redundant its contents with the radiometric information supplied by the sensor. The much used procedure of deriving basic spectral data from training sites is thus valuable only if natural crop spectra differ significantly one from the other. According to the model previously described, this occurs primarily during selected time periods corresponding to the phenologic windows. It is in these periods that training site information has maximum value.

Errors of mensuration induced by the coarseness of the pixel dimension can be reduced by appropriately combining the low-resolution radiometric data with high-resolution aerophotographic data. The latter serve to delineate geometrical boundaries, the former to identify their internal contents.

Table 6-1 presents the types of auxiliary information which most significantly contribute to the accuracy of crop inventories.

TABLE 6-1

PRINCIPAL AUXILIARY DATA PECULIAR TO REGION OF INTEREST

CROP ACREAGE INVENTORY

BASELINE DATA

o STATISTICAL DATA1 PRODUCTION, ACREAGE, PROPORTIONS OF CROPS; FIELD SIZE STATISTICS.

o PHENOLOGIC CALENDAR AND ITS STATISTICAL FLUCTUATION

o REPRESENTATIVE OR INDICATIVE SPECTRA

o METEOSTATISTICS, RAINFALL, INSOLATION, TEMPERATURE, EVAPOTRANSPIRATION

o CLOUD COVER STATISTICS

o TOPO MAPS, VARIOUS SCALES

o LATEST CONVENTIONAL AEROPHOTOS

o LATEST LAND USE MAPS

o AGRICULTURAL PRACTICES, AGRIECONOMIC DATA

o STATISTICS ON PRINCIPAL CAUSES OF CROP LOSSES

o DESCRIPTION AND DATA ON CONVENTIONAL (GROUND-BASED) INVENTORY METHODS

o CADASTRAL MAPS

DERIVED DATA

o AGROMET INDICES

o ACREAGE, YIELD, PRODUCTION TRENDS

o SIGNATURE EXTENSION INDICES

SOFTWARE

o CONVENTIONAL PROGRAMS MODULES AND PACKAGES

o SPECIAL PROGRAM CLEARINGHOUSE

DATA-RELATED SERVICES

o RECOMMENDATION OF OPTIMUM SCENE TIMING

o RECOMMENDATION OF TUNING PARAMETERS FOR SPACE-DERIVED DATA

## 7.0  SIMULATION SCHEME AND RESULTS

The approach orginally conceived was to test the theory thus far presented on a worldwide basis, including the effects of samp‌ling, phenologic cycles and cloud cover, by simulation on the Marshall Space Flight Center's Data System Dynamic Simulator (DSDS).  The DSDS concept embodies these capabilities, including capability for radiometric readout from imagery, magnetic tapes, or simulated scenes.

At the time of this simulation, the DSDS had not reached its Phase III development stage.  Hence the simulation was tailored to accept, instead of the full complement of inputs  in their orginal format, the following:  1) Synthetic sampling and discrimination derived from theory;  2)  Synthetic model of the reflectance excursions induced by phenology, derived from empirical data;  3)  Pre-determined geographic locations of the principal crops,derived from historical statistics; 4)  Synthetic cloud cover statistics, derived from existing models.

The simulation scheme was elaborated jointly with NASA Marshall and with the General Electric Company's Huntsville Operations, who ran the simulation.  General Electric's results are contained in separate report [8].  This chapter synopsizes how the simulation inputs were structured, and reports key results.  The procedure adopted was the following:

1.  The principal countries and crops were determined from the statistical material presented in Chapter 2.  Twenty-two countries were selected.  Pre-simulation analyses showed that adequate results could be obtained by simulating only the crops Winter Wheat, Spring Wheat, Corn, Soybeans, and Rice.

2. For each selected country, phenologic calendars were obtained from [9].

3. Because not all crop calendars provided complete information on all phenologic stages, they were adjusted against the data of the SRS Crop Weather Bulletin [10]. From these, U.S. phenologic statistics were compiled following the procedure described in Chapter 6, illustrated in Figure 6-3. The mean and standard deviations of each phenologic stage's time of occurrence was computed.

   Most foreign data were incomplete, e.g. provided only the mean dates of occurrence of a few phenologic stages. The phenostages were assumed to possess the same standard deviations as computed for the U.S. Subsequent work by ECOsystems [11] indicates the validity of this procedure for countries located in the temperature zones, and for tropical countries subject to regularly recurring weather phenomena, e.g. monsoons in India.

4. The phenologic phasings differed considerably among the 22 countries selected. Equivalent phenologic zones were defined as those wherein the phasing between phenostages did not exceed 20 days. Five such phenologic zones were sufficient to account for the selected 22 countries

5. The major countries, e.g. the U.S., USSR, Brazil, turned out to contain more than one phenologic zone. They were thus subdivided into sub-zones. Thirty-six regions were thus defined, as shown in Table 7-1 [8].

6. The major areas of cultivation of the five selected crops were identified and bounded in each of the 36 regions, from [14]. Analysis disclosed almost everywhere the presence of two major confusers: 1) permanent pastures, i.e. areas where grasslike crops are grown for grazing; and 2) hay-type crops, such as alfalfa, which are grasslike crops harvested

## TABLE 7-1

### COUNTRIES, REGIONS, CROPS AND NUMBER OF SAMPLES USED FOR STUDY

| COUNTRY & REGION | CROPS | SIMULATION SAMPLES | ESTIMATED NO. OF OPER-ATIONAL SAMPLE SEGMENTS |
|---|---|---|---|
| Argentina | W, C | 45 | 1500 |
| Australia | W | 30 | 1000 |
| Bangladesh | R | 30 | 1000 |
| Brazil North | C | 30 | 1000 |
| Brazil South | C, S, R | 53 | 1767 |
| Canada | W | 30 | 1000 |
| China North | W, C, S, R | 60 | 2000 |
| China Central | W, C, S, R | 60 | 2000 |
| China South | W, C, S, R | 60 | 2000 |
| Egypt | C | 30 | 1000 |
| France | W, C | 45 | 1500 |
| India Punjab | W, C | 45 | 1500 |
| India Ganges | W, C, R | 53 | 1767 |
| India Central | W, C, R | 53 | 1767 |
| India Bilaspur | W, C, R | 53 | 1767 |
| India Coastal | R | 30 | 1000 |
| Indonesia | R | 30 | 1000 |
| Italy | W, C, R | 53 | 1767 |
| Japan | R | 30 | 1000 |
| Mexico | C | 30 | 1000 |
| Pakistan | W | 30 | 1000 |
| Romania | W, C | 45 | 1500 |
| South Africa | C | 30 | 1000 |
| Philippines | R | 30 | 1000 |
| Thailand | R | 30 | 1000 |
| Turkey | W | 30 | 1000 |
| USA - Region A | W, C, S | 53 | 1500 |
| USA - Region B | W, C | 45 | 2000 |
| USA - Region C | W, C, S, R | 60 | 2000 |
| USA - Region D | W, C, S, R | 60 | 2000 |
| USSR Latvia | W, C | 45 | 1500 |
| USSR Ukraine | W, C | 45 | 1500 |
| USSR Transvolga | W, C | 45 | 1500 |
| USSR Volga-Ural | W, C | 45 | 1500 |
| USSR Siberia | W, C | 45 | 1500 |
| Yugoslavia | W, C | 45 | 1500 |
| TOTAL | | 1553 | 51355 |

W = Wheat (Winter & Spring); C = Corn; S = Soybeans; R = Rice

periodically for hay, sometimes as much as 6 to 7 times
yearly. The phenologic cycle of these two types of
confusing crops affects the discrimination of the
primary crops differently.

Permanent pastures generally reach the stage of maturity,
i.e. of high reflectance, prior to the emergence of the
wanted primary crops. It is thus necessary to "mask out"
the former from the remainder of the scene by employing
LANDSAT data gaterhed sufficiently early in the season.

The haylike crops  exhibit early season characteristics
similar to those of the pastures, yet contribute additional
information: their frequent cuttings display a characteris-
tic temporal alternation of high and low reflectances.

For reasons of simulation expediency, the haylike crops
were assimilated within the simulation model to the
permanent pastures. A corresponding pattern of reflectance
progression was constructed for each of the 36 regions.

7. The phenologic "windows" were constructed for each of the 36
   regions. The mean reflectances corresponding to each pheno-
   stage were assumed for all regions to equal those observed
   in the available U.S. experimental data. The fluctuations
   around the means were assumed to equal $\pm 30\%$ one sigma, and
   to be uncorrelated among different crops. The unperturbed
   threshold of separation between reflectances was set at
   in 20 years. This means that deviations from the reflectance
   of two crops must have a joint probability of occurrence of
   0.05, i.e. individual probabilities of occurrence of $(0.05)^{\frac{1}{2}} =$
   0.22. On the assumption of normal distribution of the
   fluctuations, this yields a threshold fluctuation level of
   0.76 for each of the two crops to be discriminated or 23%
   of their mean reflectances. The statistical window durations
   were set to the 95% confidence level: i.e. windows shorter
   than those which were modeled occur on the average less
   frequently than every 20 years.

8.  An average sensitivity threshold was computed for all win-
    dows, by RSS combinations of :  J) 1% for sensor fluctua-
    tions;  2) + 6% for atmospheric fluctuations as described
    in Chpater 6.  In order not to overly complicate the
    simulation, the sensitivity threshold was assumed fixed
    rather than proporticnal to the reflectance's amplitude:
    its value was selected as 0.02, or 2% of the relative
    maximum reflectance of 1. Typical examples of windows
    are depicted in Table 7-2.

TABLE 7-2

EXAMPLES OF WINDOW FORMAT USED IN SIMULATION

| COUNTRY | REGION | CROPS | WINDOW STARTS JULIAN DATE | WINDOW DURATIONS DAYS |
|---|---|---|---|---|
| U.S. | A (IL, IO, OK, KS) | WW-P | 111 | 27 |
| | B (MT, SD, ND, UT, CO) | SW-P | 118 | 40 |
| | C (VA, KY, NC, TN, SC, CA, AL, MI, LO) | C-S | 151 | 17 |
| USSR | UKRAINE | WW-S | 115 | 48 |

WW = Winter Wheat, SW = Spring Wheat, P = Pasture, C = Corn
S = Soybeans

9.  A sampling scheme was constructed within each region.  Based
    upon the results derived in Chapter 6, good stratification
    and bounding was assumed:  each candidate region was assigned
    1,000 sample segments if it contained a single crop:  and
    1,500; 1767; 2000 segments if it contained two three, four crops
    respectively. For expediency of simulation, sub-multiples or
    these numbers of sample segments were chosen as inputs to
    DSDS:  specifically, 30, 45, 53 and 60 for single, double,

triple and quadruple crops respectively. This reduced set
of segments was selected at random from the larger set. The
list of countries, regions, crops and number of sample seg-
ments contained within each is shown in Table 7-1. Not there-
in shown but implicit in the previous construction of the
windows is the presence of the selected confuser, pasture,
discussed under point 7, preceding.

10. Simulation of the complete logic of the cumulation of informa-
    tion from multiple looks exceeded the constraints placed on
    this effort. In its place, target acquisition was specified
    as occurring when at least one "look" per phenologic window
    was possible. Whenever such look did not occur, either due
    to excessive shortness of the windows or impeded by the pre-
    sence of cloud cover, the particular region and crop affected
    was considered as not having been surveyed.

    The "one look" assumption, necessitated by simulation expe-
    diency, is reasonably adequate for the purposes of this pro-
    ject, because:

    a. The specification of a single look per window is some-
       what lax if the single look is all that is available.
       ECOsystems experience with real scenes shows that
       generally more frequent looks are needed: an appro-
       ximate average is of order 1.5 per window.

    b. The requirement of one look per window is somewhat
       stringent if a sufficient number of earlier or later
       looks are provided wherefrom to reconstruct the state
       of affairs within the missed window.

11. Cloud cover statistics were incorporated in the simulation
    by the General Electric Company, in accordance with the available
    worldwide Allied cloud model. The model's statistics are sup-
    plied in terms of percent cloud cover.

A more cogent statistic, i.e., the "probability of effecting target acquisition," requires knowledge of the statistics of unobstructed vision within each scene, i.e. of the patterns of "holes in the clouds." The DSDS, in its full phase III implementation, would have been capable of deriving these statistics from automated analysis of historical LANDSAT and/or GOES and NIMBUS scenes. In its absence, it was instead decided to incorporate into the simulation two alternate cloud cover thresholds: 50% and 90%. When cloud cover exceeded the threshold, the acquisition of a crop target was considered missed.

The principal results derived from the simulation for the LANDSAT-D configuration are indicated in Table 7-3.

TABLE 7-3

PERCENT OF SIMULATED SAMPLE SEGMENTS ACQUIRED BY A SINGLE
ORBITING PLATFORM [ 8 ]

PLATFORM CHARACTERISTICS:
    ORBITAL ALTITUDE:  704 km
    ORBIT INCLINATION:  SUN-SYNCHRONOUS
    SWATH WIDTH:  185 km

| COUNTRY/REGION | PERCENT @ CLOUD COVER 50% | @ CLOUD COVER 90% |
|---|---|---|
| ARGENTINA | 93.3 | 97.8 |
| AUSTRALIA | 95.0 | 100.0 |
| BANGLADESH | 84.4 | 95.6 |
| BRAZIL - NORTH | 100.0 | 96.7 |
| BRAZIL - SOUTH | 99.1 | 98.1 |
| CANADA | 96.7 | 96.7 |
| CHINA - NORTH | 85.8 | 88.3 |
| CHINA - CENTRAL | 76.7 | 90.8 |
| CHINA - SOUTH | 87.5 | 90.0 |
| EGYPT | 100.0 | 100.0 |
| FRANCE | 86.7 | 97.8 |
| INDIA - PUNJAB | 87.8 | 98.9 |
| INDIA - GANGES | 75.0 | 85.4 |
| INDIA - CENTRAL | 72.3 | 90.6 |
| INDIA - BILASPUR | 82.4 | 89.9 |
| INDIA - COASTAL | 85.6 | 87.8 |
| INDONESIA | 85.0 | 91.1 |
| ITALY | 86.9 | 93.5 |
| JAPAN | 95.0 | 93.3 |
| MEXICO | 97.2 | 98.9 |
| PAKISTAN | 93.3 | 80.0 |
| ROMANIA | 91.1 | 91.1 |
| S. AFRICA | 90.0 | 93.3 |
| PHILIPPINES | 91.7 | 96.7 |
| THAILAND | 91.7 | 96.7 |
| TURKEY | 93.3 | 80.0 |
| US - A | 80.0 | 80.0 |
| US - B | 93.0 | 93.0 |
| US - C | 84.6 | 87.9 |
| US - D | 77.0 | 84.7 |
| USSR - LATVIA | 96.7 | 100.0 |
| USSR - UKRAINE | 96.7 | 96.7 |
| USSR - TRANS-VOLGA | 97.8 | 95.6 |
| USSR - VOLGA-URAL | 98.3 | 96.7 |
| USSR - SIBERIA | 94.4 | 97.8 |
| YUGOSLAVIA | 86.7 | 95.6 |
| WORLD AVERAGE | 89.1 | 92.4 |

C-4

The Table shows results similar to those presented in previous Figure 6-5: namely, a single orbiting platform with the characteristics of LANDSAT-D achieves order of 90% of the desired worldwide observations. Whereas this result might at first blush appear incommensurate with the specified accuracy of 96-98% worldwide, it does not tell the entire story: for not all countries and regions missed contribute equally to the world's production of the identified four staple crops. A more sophisticated simulation scheme is required to determine the weighted impact of missed observations.

In any event, modifications to future earth observation systems, other than the brute force expedient of multiple spacecraft, appear desirable and practically feasible, as discussed in the following chapter.

## 8.0  IMPACT UPON OSTA'S AGRICULTURAL DATA SYSTEMS OF THE NEXT DECADE

### 8.1  State of the Data System

The reason for being of OSTA Space Systems is to provide information. Thus the Data System represents their paramount element, which all other Subsystems are designed to support.

To compete with other acreage-measurement systems - e.g. aircraft, ground surveys - OSTA Agricultural Data Systems must provide better price/ performance. Industry tests the price/performance and market acceptance of a given product, prior to its entry to market, by a process known as product assurance. A fallout from this study is a preliminary product assurance review of OSTA agricultural information systems.

The principal design recommendations which affect the overall system's price-performance in Agricultural Applications are recapitulated following.

### 8.2  System Performance - Accuracy

To achieve market acceptance, the Data System's accuracy need be increased from the currently achieved average level of approximately 80-85%, one sigma, to 96-98%, i.e. by a factor of at least three, preferably five. Except for the interference from cloud cover, there are no theoretical reasons barring this achievement. Improvements of on-board radiometric calibration, better registration of images, and similar sensor-oriented procedures, will assist but will not correct the problem.        principal

bottleneck lies in the information-extraction subsytem: specifically it is attributable to the manner in which radiometric information has been heretofore exploited.

Prinicpal corrective actions are:

● Exploitation of the phenologic windows phenomenon

● Integration of auxiliary information with the radiometric sensing data.

The first requires significant increases of the observations's recurrence frequency. The second requires making available suitable external information, approximately formatted.

### 8.2.1 System Performance-Recurrence Frequency

The data recurrence frequency must be sufficiently high to cope with the finite time duration of the temporal windows, further reduced by cloud cover.

The polar-orbit, sun-synchronous, 185 km swath configuration characteristic of LANDSAT supplies adequate worldwide crop visibility if multiple (order of three), suitably phased orbiting platforms are provided.

Promising avenues to obviate this costly expedient are:

● Use of microwave sensing in conjunction with optical and near-IR sensing

The technical feasibility of active microwave sensinf from orbit was tested in the SEASAT-A spacecraft.

Limited experimental data in crop reflectances in the microwave spectrum indicate strong analogies with their behavior at optical wavelengths. Theory supports these findings.

- **Increase in the sensor's swath width**

  Current thinking holds that nadir viewing of agricultural crops is preferable to observations from slant angles. No adequate experimental verification of this assumption was found in the literature: in fact, limited experimental data [ ] point to advantages from slant observations.

  The effect of increased swath widths is to lower the repetition interval required of the observations. Its drawback, i.e. reduced cross-track resolution or the need of more sophisticated sensors to maintain resolution, can be obviated by replacing wall-to-wall with sampled observations.

- **Modification of orbital parameters.**

  Reductions of recurrence interval are achievable by modifying the platforms orbit. Those are however of lesser magnitude than what can be accomplished by the methods previously discussed.

  The largest effect occurs if the orbital inclination is reduced so as to cover only the significant agricultural and/or forested regions of the world. Figure 8-1 suggests an orbital inclination of order $60^{\circ}$.

  This system modification would entail a departure from sun-synchronous operation. The corresponding effects upon discrimination remain to be tested experimentally. It is noted that no adequate justification for sun-synchronism was found in the literature: numerous discussions with members of the remote sensing community' indicate that the specification of this requirement is at present founded essentially upon expert opinion rather than upon observed fact.

FIGURE 8-1

CUMULATIVE VALUE OF WORLD'S AGRICULTURAL AND FORESTRY PRODUCTS VERSUS LATITIDE

LATITIUDE - DEGREES
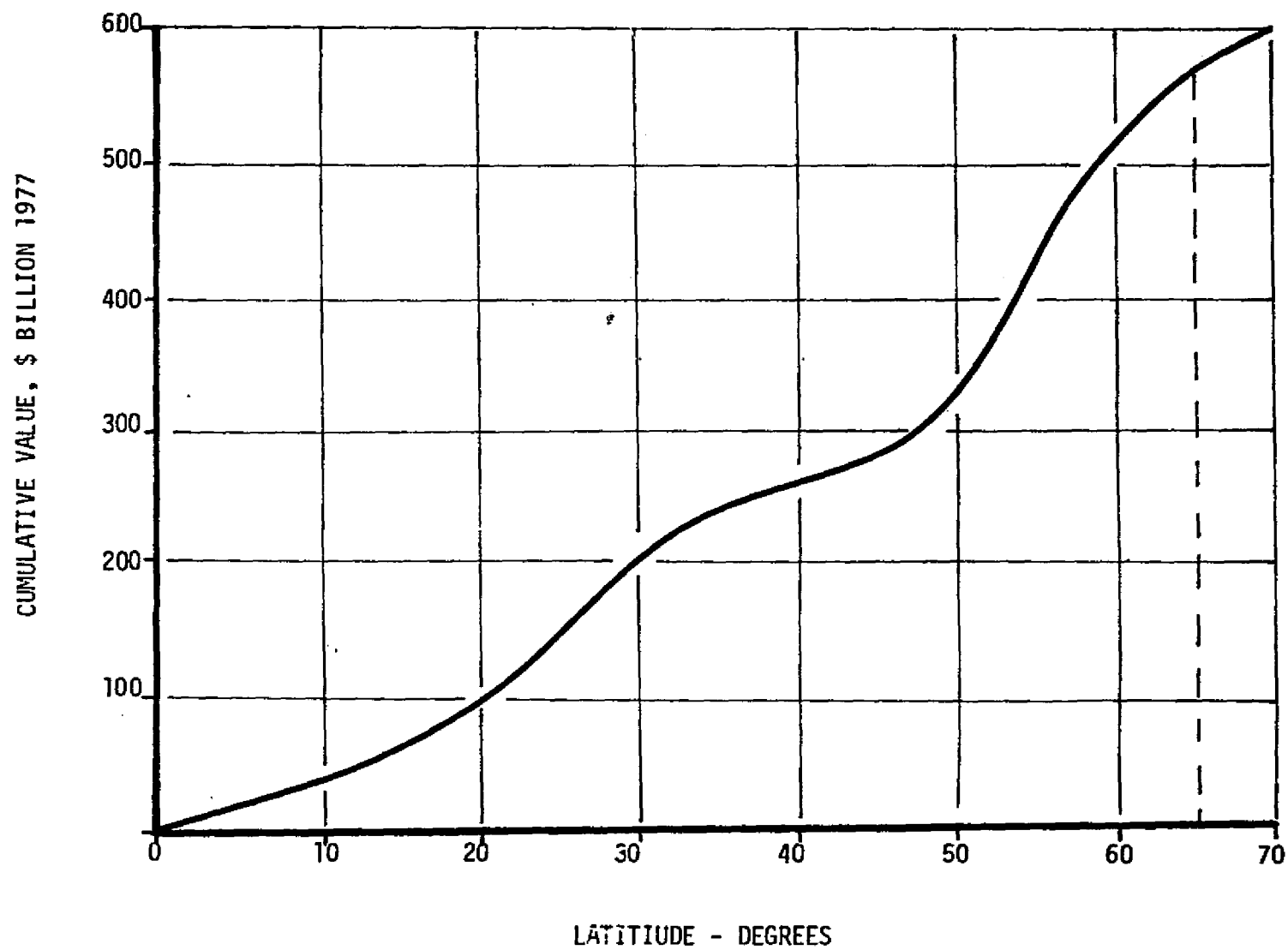
## 8.2.2  System Performance - Role of Auxiliary Information

It is necessary to combine radiometric imagery with auxiliary information to reach levels of performance comptaible with the user's requirements.  Most US auxiliary data are avialble frcm data bases managed by Federal Agencies; the residual is obtainable at State and Local levels.

The primary sources of foreign data are agencies at the Ministerial level.  Albeit with exceptions, foreign data are significanly scantier and of more difficult access than in the US.  The situation particulary critical with regard to aerophotography; even where available, this most important auxiliary information is so time-consuming and laborious to retrieve as to very seriously affect the final product's timelines and price.

This situation can be obviated by providing easily accessible global photographic base coverage, at appropriate geometric resolutions. The refresh cycle should be commensurate with the period of obsolence of the observables: in the case of agriculture, it need be no shorter than two or three years.

A means of implementing this service would be to reduce the RBV'S field of view, coupled with progressively stepped side look.  Table 8-1 shows the times-resolution tradeoffs for complete coverage of the earth surface.

TABLE 8-1

TIME REQUIRED FOR COMPLETE RBV COVERAGE OF THE EARTH'S
SURFACE VERSUS RESOLUTION

| RESOLUTION, METERS | COVERAGE TIME REQUIRED, YEARS |
|---|---|
| 40 | 0.2 |
| 20 | 0.7 |
| 10 | 2.85 |

It can be seen that resolutions adequate for agricultural appli-
cations are achievable within practical time constraints.

## 8.3 System Costs

The wall-to-wall method of data acquisition is costly on two accounts:

o  It requires very high data relay rates (84 Mbit/sec for LANDSAT D)

o  It requires pre-processing subsystems of high throughput

The peak throughput requirements are particularly high in certain
periods of time.  During the months from March through August in the
northern hemisphere, and their counterparts in the southern, the simu-
lation shows that as much as 500 observations are required daily.

Preliminary computations [    ] indicate that as much as 50%
of the LANDSAT D system's running costs is attributable to the com-
bination of TDRS channel rental lus cost of pre-processing.

As shown in Figure 6-11, reduction of the peak data load by as much as 30:1 is achievable by employing sampling as the primary data-gathering mode. Wall-to-wall data are required to provide {ase information, but only occasionally.

The presence of reflectance fluctuations indicates another possible avenue of price/performance enhancement for agricultural observation systems, through relaxation of the sensor's NEΔ specifications. The sensor's sensitivity should be matched to the overall system's error budget, of which the "scene fluctuation" is the major contributor.

Sensor sensitivities of order 64 levels appear adequate. The resulting 4.1 reduction of sensor sensitivity can be traded off against lower-cost sensor implementation, or it can be exploited to increase the sensor's geometric resolution. It can be shown theoretically that the complete tradeoff would improve LANDSAT'D resolution to approximately 15 meters.

# APPENDICES

APPENDIX A    THE UN/FAO SCENARIO FOR GLOBAL CROP FORECASTING


APPENDIX B    CURRENT CONVENTIONAL METHODS OF AGRICULTURAL CROP
              MEASUREMENT AND FORECASTING


APPENDIX C    DETERMINATION OF THE ACCURACIES OF AGRICULTURAL ACREAGE
              ESTIMATION ACHIEVED BY ORBITING DATA SYSTEMS


APPENDIX D    ANALYSIS OF AVAILABLE CROP SPECTRA


APPENDIX E    INFORMATION EXTRACTION FROM RADIOMETRIC DATA
              DISCRIMINATION THEORY

A-1

# APPENDIX A

## THE UN/FAO SCENARIO FOR GLOBAL CROP FORECASTING

The UN/FAO viewpoint can be synthesized by the following logical
sequence of propositions:

1.  The World's human food supply is dominated by cereals, which
    account for 53% of the total, followed by 12% tubers, 15%
    high-protein foods (meat, fish, milk), 20% vegetables, fats,
    sugar combined [2].

2.  The global growth of cereal production outstrips population
    growth. Figure A-1 appears to indicate that no global nutrition
    problem exists [12].

3.  However, more detailed analysis of the cereal production and popu-
    lation trends of each of the three FAO zones, as in Figure A-2, re-
    veals that the problem is one of distribution: in the most populous
    Zone C, production increases barely keep pace with population growth.
    Because Zone C food intake is now only approximately 80% of the FAO
    norm, Zone C appears to be hopelessly undernourished.

4.  A deeper step in the analysis, portrayed in Figure A-3,
    indicates that even in Zone C, cereal availability per
    capita significantly outstrips the consumption, even though the
    apparent per capita consumption is quite high.

5.  What causes the apparent inconsistency? Figure A-4 provides the
    explanation. Almost three billion large farm animals, not counting
    upwards of 10 billion heads of poultry, share the earth's surface
    with man. Up until the late fifties, these animals subsisted pri-
    marily on pasture and grass products; in the last decade and a
    half, the composition of feed has been significantly enriched by

# FIGURE A-1

## THE GLOBAL FOOD OUTLOOK



PERCENT HUMAN CONSUMPTION

100% — SUGAR / FATS & OILS / VEGETABLES & FRUITS / TUBERS

75% — MEAT, FISH, MILK

50%

25% — CEREALS

HUMAN FOOD CONSUMPTION IS PRIMARILY BASED ON CEREALS.

PERCENT GROWTH



150%
140%
130% — CEREAL PRODUCTION
120%
110% — POPULATION
100%
61 62 63 64 65 66 67 68 69 70 71 72 73

YEARS
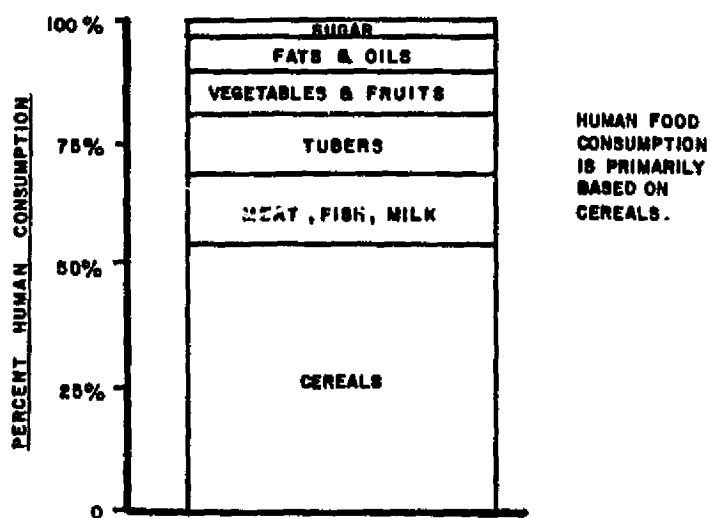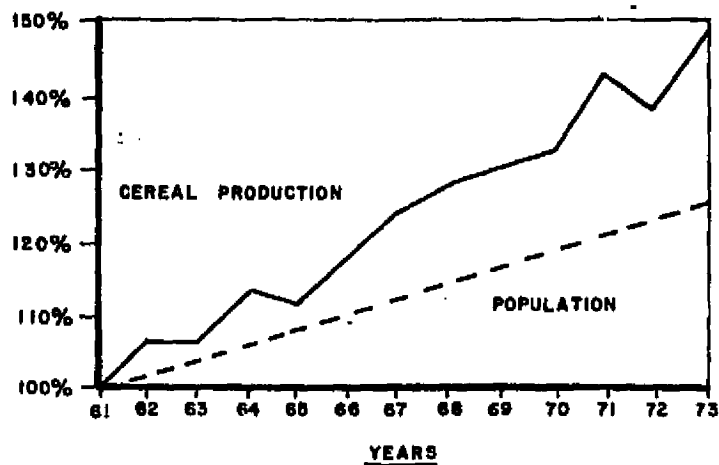
# FIGURE A-2

## THE THREE FAO WORLD ZONES



A  Developed Market Economies
B  Planned Market Economies
C  Developing
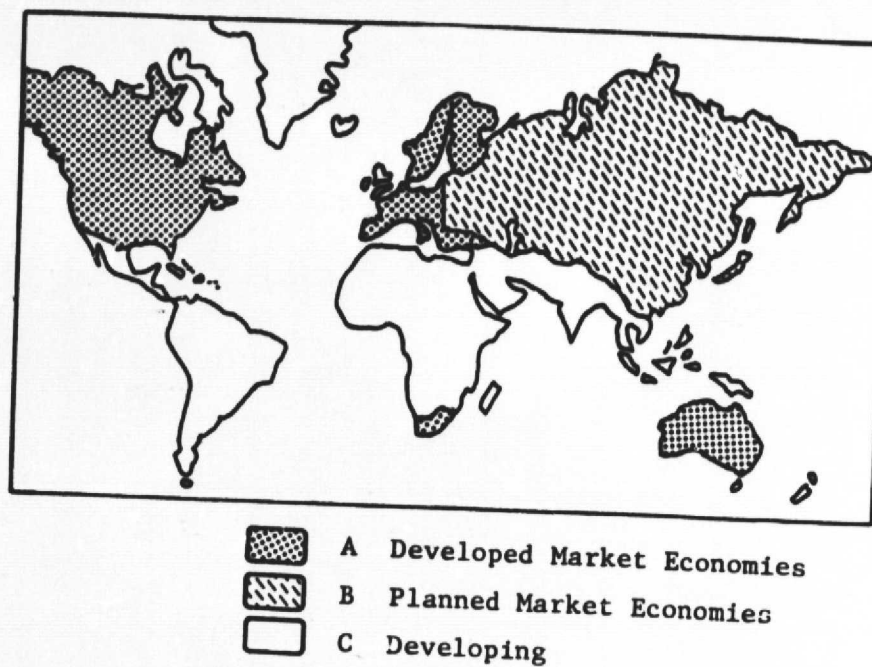
## THE ZONAL FOOD OUTLOOK
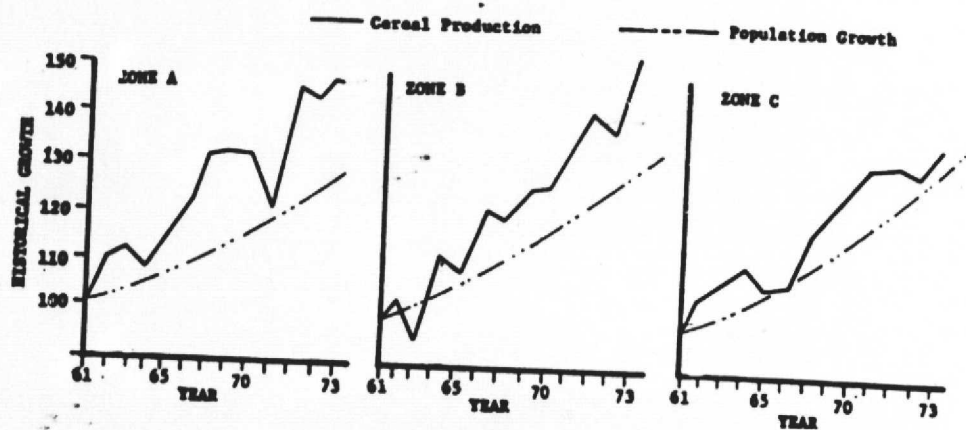
## FIGURE A-3

## THE WORLD CEREAL BUDGET BY FAO ZONES

### AVAILABILITY OF CEREALS



| POPULATION billions | PRODUCTION megatons | PER CAPITA kg/yr. |
|---|---|---|
| 1.8 | 420 | 231 |
| 1.24 | 395 | 315 |
| 0.96 | 325 | 338 |
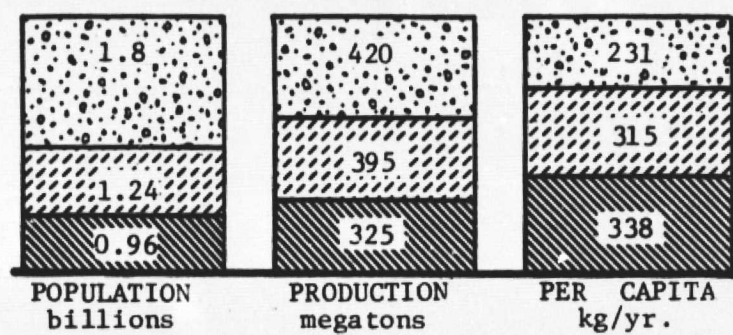
### HUMAN CONSUMPTION VERSUS AVAILABILITY



Kg/Capita/yr.

Zone A   Zone B   Zone C

ZONE A

ZONE B

ZONE C

consumption

discretionary surplus
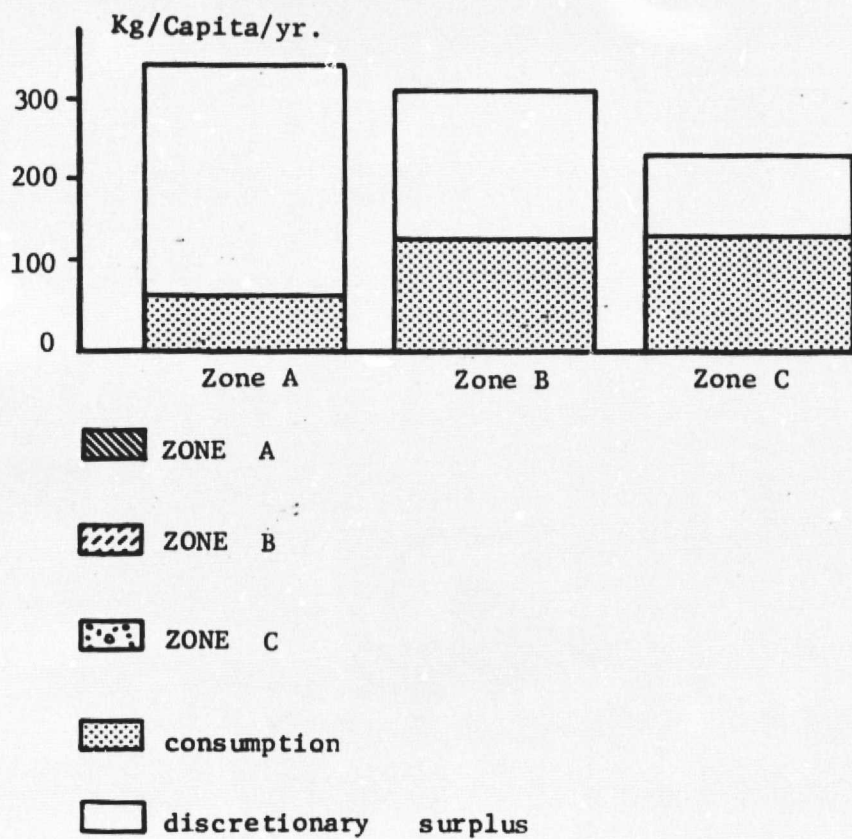
# FIGURE A-4

## THE INCREASING USE OF CEREAL IN ANIMAL FEED



CURRENT WORLD MEAT SUPPLY

BOVINES 956  OVINES 1066  SUINES 634.5  Production 92.5 MT  PER CAPITA Kg/yr.

12.5  25.7  39

Zone A   Zone B   Zone C

HISTORICAL TREND IN ANIMAL FEED

Megatons

HAY

CEREALS

100

50

1962   1968   1974   U.S.

Megatons

1200
1000
800
600
400
200
0

SUCCULENT
SILAGE
COARSE HAY
PASTURE

1965   1970   1973   U.S.S.R.
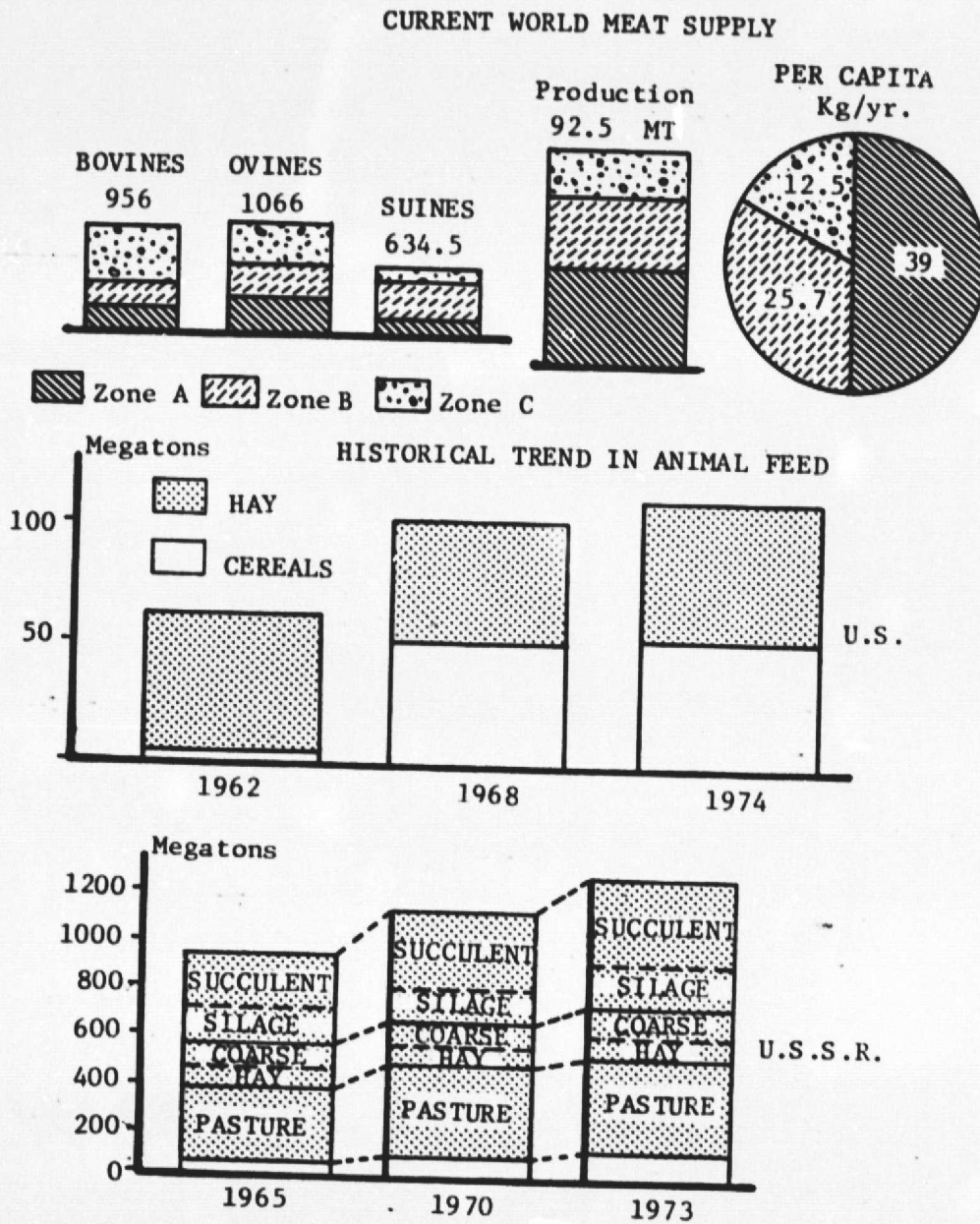
the so-called "concentrates," i.e. cereal products: corn, soy, wheat.

6. The reason? The growing demand for protein: its efficient and cost/effective production requires concentrated animal fodder. This "technological revolution" has vastly increased the importance of cereal products, much beyond the proportions depicted in Figure A-1 for human consumption alone. Since meat, milk and animal fat production is now driven by cereals availability, it is safe to estimate that circa 75% of global human food consumption is based on cereals.

7. The cure? Figure A-5 indicates that on the average, Zone A's surplus just matches Zone C's deficit. The simple device of straightforward transfer would thus, at least on the surface, resolve the global food problem. That is, except for the economic implications, i.e. the matter of who would foot the bill.

8. The cost of such transfer was computed by the FAO in their recently compiled World Indicative Plan. As plotted in Figure A-5 the straightforward transfer of food including acquisition and shipment, would cost $63 billion per year, growing to approximately $80 billion in 1985. The cost of financing adequate infrastructures -- fertilizer plants, farm machinery, education, etc. -- to enable Zone C to become adequately self-sufficient would cost $25 billion yearly growing to approximately $30 billion a decade hence. In practice, a 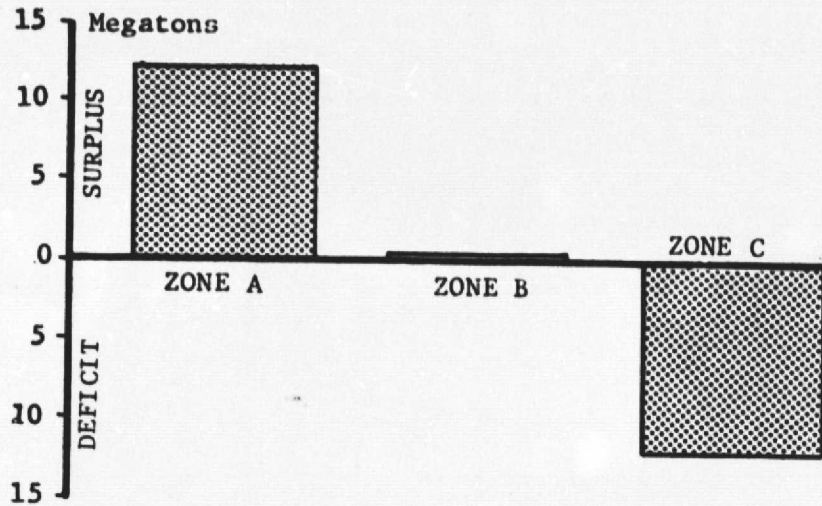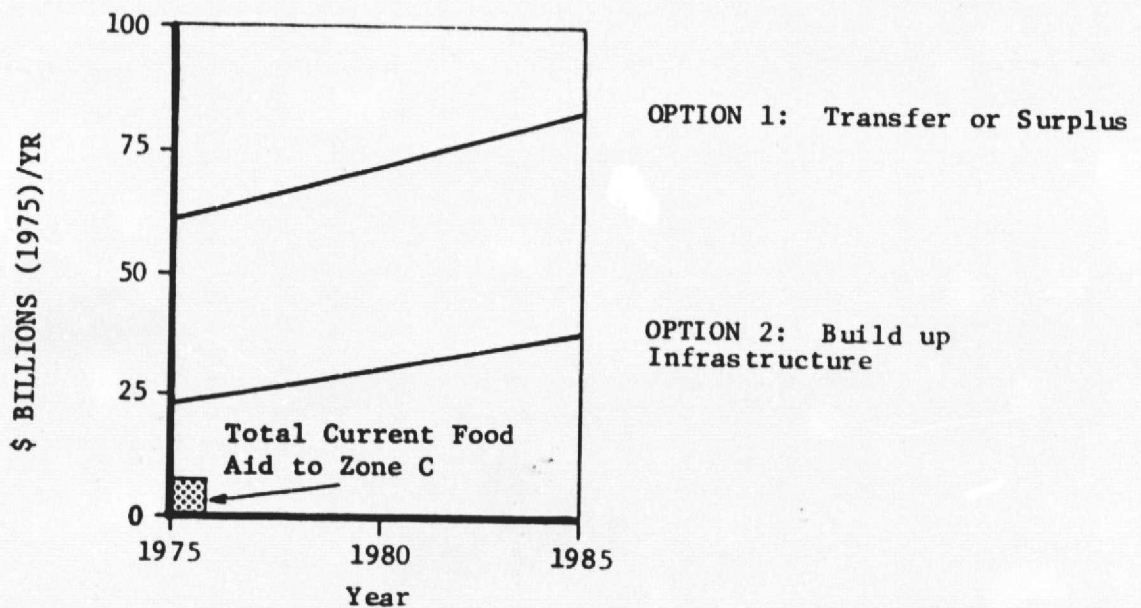complete solution of the world's food problem would consist of a mixture of these two methods. In the limit, it would entail a yearly expenditure of almost $90 billion yearly currently, growing to exceed $100 billion yearly in the mid-1980's.

9. While these expenditures are not impossibly high -- the U.S. share, for example, of the "maximum plan" would not exceed $36 billion yearly now, growing to $45 billion in 1985, or approximately one-third of our current National Defense outaly -- they are not likely to be implemented in the near future.

# FIGURE A-5

## THE WORLD FOOD DISTRIBUTION PROBLEM

### YEARLY CEREAL SURPLUS AND DEFICIT
### 10 year average (1963-1973)



### COST OF TOTAL FOOD
### ASSISTANCE TO ZONE C



OPTION 1:　Transfer or Surplus

OPTION 2:　Build up Infrastructure

Total Current Food Aid to Zone C

Note that the contributions for agricultural aid to Zone C in all forms, from all of Zone A and B combined, cover not more than approximately 10% of the requirements stated by the FAO.

The analysis above leads to the conclusion that the primary requirement to alleviate the global food problem is to manage as best as possible the scarce amount of agricultural aid currently flowing from Areas A and B into Area C.

The requirements "universe" as seen from the UN/FAO outlook can be divided logically as follows:

1. Countries in which internal production consistently equals the demand, and which are thus self-sufficient nutritionally.

2. Countries whose internal production is consistently less than the demand, but which are able to import the deficiency through other sources of revenue. Typically, the OPEC countries.

3. Countries with consistent or periodic production-demand deficits and which are unable to fill the deficiency by imports.

4. Countries which are consistent overproducers and therefore exporters.

Countries which exhibit characteristic 1 are not important to the UN/FAO goal, because they already meet it: nor can they directly assist in transferring surplus to other diet-deficient Countries.

Countries exhibiting characteristic 2, i.e. which are consistently in crop deficit but which can import the deficit through other means of revenue, are only important to the extent that their fluctuating import requirements may increase or reduce the surplus available to the chronically deficient Countries.

Survey of Countries in Category 3, chronically deficient and unable to import the deficit, are of interest to assess the quantity of product they are likely to have available in any one year, and thus the extent of the current deficiency. This assessment serves as the basis for UN/FAO is planning of the amount of food aid to be provided. Aside from political considerations, the interest is proportional to the extent of the deficit and the number of persons affected.

Surveys of Countries in Category 4, consistent overproducers, are of primary interest in proportion to the extent of discretionary surplus they historically produce. The assessment of their production serves as the basis for planning the most appropriate distribution of food aid.

It is worth noting that the UN/FAO scenario presented above is idealized: in current practice, the food aid funds at the disposal of the UN are insufficient to cope with the demand. Also, AID-dispensing Countries do not act in harmony and nor for the solely humanitarian purpose of providing their surplus where most needed: rather, political considerations tend to channel food aid where each donor deems it to be most opportune.

In spite of these departures from ideal perfection, the UN/FAO scenario is of value as an approximate model to guide the equalization between deficits and surpluses.

Table A-1 depicts the aggregate distribution of human consumption in diet-deficient regions. It shows the presence of the staples already identified in the analysis of the U.S. balance of trade situation, plus the emergence of rice as a new major staple.

Table A-1 also indicates the 6-year cumulative import-export trends for rice, and its principal producing Countries.

Table A-2 tabulates the Countries and crops candidates for survey as seen from the UN/FAO outlook.

## APPENDIX B

### CURRENT CONVENTIONAL METHODS OF AGRICULTURAL CROP

### MEASUREMENT AND FORECASTING

The measurement and forecasting of agricultural production are activities of major importance in virtually all countries of the world, as indicated in Figure B-1 [13].

The most sophisticated forecast system is the United States Department of Agriculture (USDA's) Economic and Commodity Reporting Service (ERCS)  Since much information is available on this system, it can conveniently be used to exemplify conventional procedures.

The key functions of the ERCS data gathering, information development and dissemination cycle are shown schematically in Figure B-2 [1].

Shown in Figure B-3 are SRS's three principal sources of data: a)  voluntary reports from individual farmers (9.5 million question-naires mailed yearly, approximately 3 million responses); b)  objective yield measurements by specialists; c)  direct sampling of selected farms by specialists (enumerators), through farmer interviews supported by map and aerial survey information.

A typical reporting form periodically filled out by individual farmers is shown in Figure B-4.  Aerial photography at scale of order 1:20,000 is used by USDA's enumerators to precisely mensurate a limited number (approximately 17,000, covering 0.6% of total farm
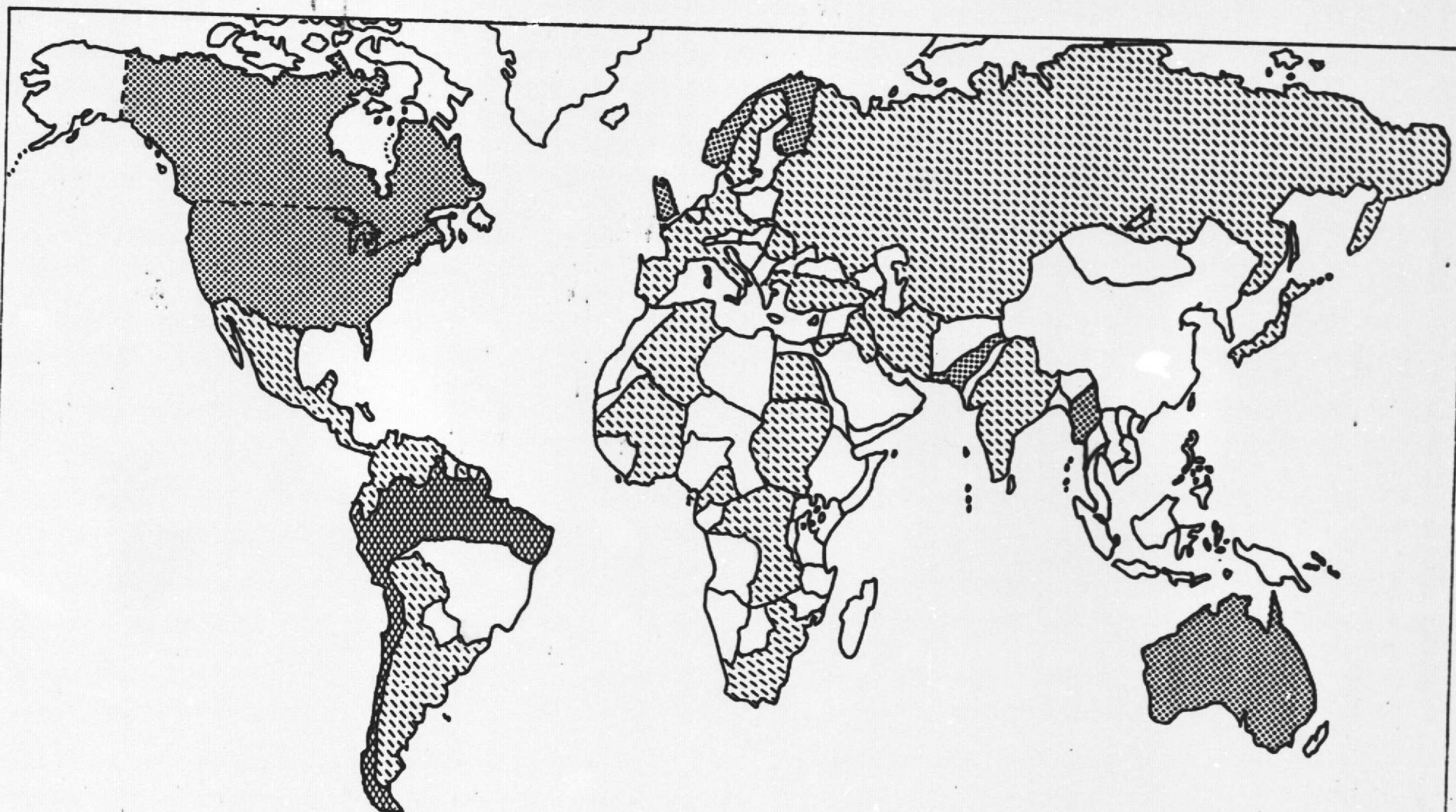
Agricultural Data Gathering System

▓ Sophisticated     ▨ Medium Complexity     ▦ Simple     ☐ No Data

FIGURE B-1
Distribution of Crop Forecasting Activities

## FIGURE B-2

### TYPICAL USDA (CORN, WHEAT) FORECAST AND FORECAST RELEASE CYCLE

REPORTS:

| MARCH CROP PRODUCTION PROSPECTIVE PLANTING | MAY WHEAT SITUATION | HIGHLIGHTS OF U.S. CROP REPORT — AUG | CROP PRODUCTION ANNUAL SUMMARY ACREAGE YIELD | AGRICULTURE STATISTICS 19__ |

PLANTING INTENTION  
PRODUCTION OUTLOOK  
CONDITION / YIELD  
FINAL ESTIMATE  
REVISED ESTIMATE

J F M A M J J A S O N D D FOLLOWING 2 YEARS

▼ = Reports to Public

□ = Forecasts
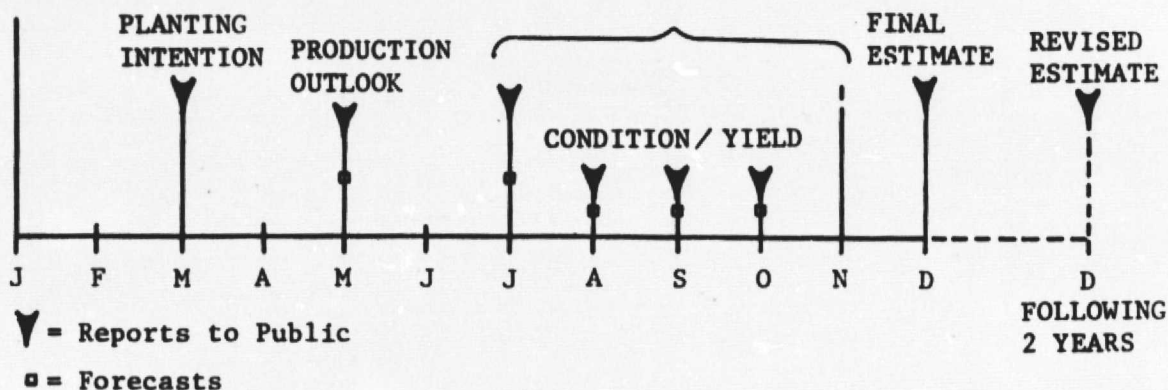
## FIGURE B-3

### SIMPLIFIED USDA DATA GATHERING PROCEDURE

44 State Offices

U.S.D.A. Headquarters

OBJECTIVE YIELD PROGRAM

STATE FORECAST

CROP REPORTING BOARD

FORECAST TO PUBLIC

700 YIELD SPECIALISTS 10,000 SAMPLES

9.5 MILLION QUERIES

3 MILLION REPLIES

ENUMERATOR PROGRAM

ACREAGE PRODUCTION

INTENT ACREAGE CONDITION YIELD

16843 SEGMENTS 2 TO 5 FARMS PER SEGMENT

1500 ENUMERATORS

AREOPHOTOS

SRS UTILIZATION 0.6% OF U.S. AREA

ASCS AEROSURVEY PROGRAM

100 MILLION HA PER YR

area) of agricultural producing units. The farmer's information is validated by planimetering the farm's area, accurately identified on the photograph.

Yield prediction (bushels per acre or quintals per hectare) is performed by: a) the farmer, who assesses the "condition" of the crop (its status with respect to what it was at the same time last year); and b) specialized personnel who perform "objective measurements" of various indicators of plant development on selected test plots; e.g., plant density, number of ears, number of spikelets, and so forth.

These reports are integrated by the central Crop Reporting Board, which issues monthly forecasts beginning 4 to 6 months prior to harvest, depending upon the crop. A final estimate of production for the current year is given shortly after the harvest.

USDA corrects their estimates over the following two years, based on additional information which becomes available on the amounts of grain products actually processed and sold by food industries. As evidenced by Figure B-5, the corrections are quite small when compared with the variances of selected other countries. Figure B-5 indicates that the fluctuations in typical developing Nations reflect a considerably higher degree of uncertainty than exists in developed Nations.

USDA's method of forecasting production is based upon the product of two factors: 1) the acreage planted; and 2) the estimated yield (bushels per acre or quintals per hectare). The yield estimate is based upon a regression formula of the type:

| CROP | ACRES FOR HARVEST | |
|---|---|---|
| | THIS YR | LAST YR |
| 1. CORN | | |
| 2. WHEAT | | |
| 3. OATS | | |
| 4. BARLEY | | |
| 5. RYE | | |

ACRES OF ALL LAND IN THIS FARM
(INCLUDE LAND RENTED)

NAME  _ _ _ _ _ _ _ _ _ _ _ _

ADDRESS _ _ _ _ _ _ _ _ _ _ _ _

DEPICT AT THIS TIME. LET 100%
REPRESENT A NORMAL CONDITION

| JULY 1 | ANSWER HERE |
|---|---|
| REPORT FOR YOUR FARM: | |
| ACRES OF CORN HARVESTED LAST YEAR _ _ _ _ _ ACRES | _ _ _ _ _ _ |
| ACRES FOR HARVEST THIS YEAR | _ _ _ _ _ _ |
| REPORT FOR YOUR LOCALITY | |
| ACRES FOR HARVEST THIS YEAR COMPARED | _ _ _ |

PLEASE ANSWER THESE QUESTIONS
FOR THE FARM YOU OPERATE

CROP PRODUCTION & STOCK

| | |
|---|---|
| CORN PRODUCED ON THIS FARM LAST YR 70 LB. EAR OR 56 LB. SHELLED BUSHEL | |
| CORN ON THIS FARM JAN. 1, 1974 70 LB. EAR OR 56 LB. SHELLED BUSHEL | |
| ALL WHEAT PRODUCED ON THIS FARM LAST YEAR  60 LB. BUSHELS | |
| ALL WHEAT ON THIS FARM JAN. 1, 1974 60 LB. BUSHELS | |

FIGURE  B-4

TYPICAL REPORT FORMATS, U.S. FARMERS
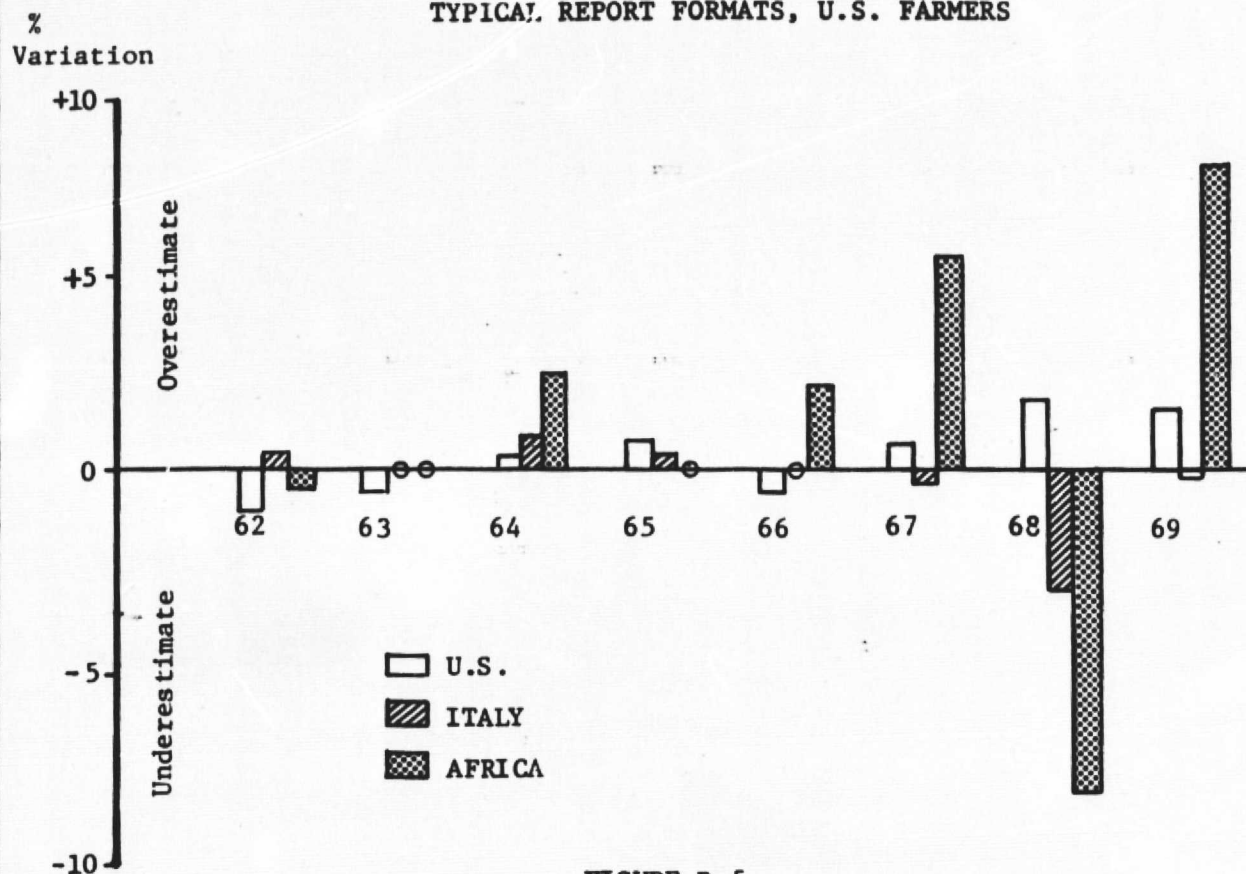


FIGURE B-5

VARIATION OF FINAL ESTIMATE

$$Y = aX_1 + bX_2 + cX_3 + dX_4$$

where:

$X_1$ is the "condition" of the crop at time of estimate, reported by the farmer or crop surveyor;

$X_2$ is the precipitation, in inches, which occurred during the previous two months;

$X_3$ is the precipitation, in inches, which is predicted will fall in the following two months;

$X_4$ is the historical growth of yield: for wheat, for example, approximately 3% per annum.

a, b, c, d are coefficients whose value is derived from regressing historical records over approximately the previous fifteen years.

Implicit in this formula is the hypothesis that the methods of cultivation: a) fertilizer input; b) weeding effectiveness; c) density of planting, and so forth, are essentially constant from year to year, and will be practiced by the individual farmer with the same diligence and with the best technology known at present. Also implicit is the assumption of a "normal" year as regards crop disease. The occurrence of diseases such as wheat rust or corn blight, or unusual events such as floods, is factored in by the farmers and specialized survey personnel, who amend their estimate, upon detection, in the next monthly reporting period.

Note that in the formulation above, the only corrector of the estimates is precipitation: rainfall is assumed to be the principal

driving function for yield. While this assumption has been found adequate for the U.S., it may not be optimal for Nations subject to different climatologies. France, for example, employs the additional variable of integrated temperature (degree-days) for their yield model.

A measure of the accuracy of forecasts referenced to the "final" yearly estimate for the last several years is indicated in Figure B-6. Figure B-7 depicts the average historical trend of forecast accuracies. Figure B-8 shows the average accuracy experienced over the last 15 years for winter wheat, the major variety of U.S. wheat.

The relative contribution to the forecast of the acreage and yield components is shown in Figure B-9, which compares the fluctuations in U.S. wheat production with the variations in acreage and yield. Note that acreage fluctuations account for almost 50% of the total production variance.

The USDA system is substantially followed, with varying degrees of sophistication, by the developed Nations. In developing Nations, simpler systems are employed. Table B-1, based upon FAO data, categorizes the crop reporting systems employed by the majority of the world's nations. Note that the complete system composed of farmer reports, plus objective yield measurement plus enumeration by specialists is used in general in those countries which possess the highest levels of living and literacy rate. As the GNP per capita diminishes, crop measurement systems tend to eliminate farmer reports and rely primarily upon sample measurements by specialists. The frequency and sophistication of the sampling procedures

FIGURE B-7

TREND OF AVERAGE USDA FORECASTING ERRORS

(Major Commodities)



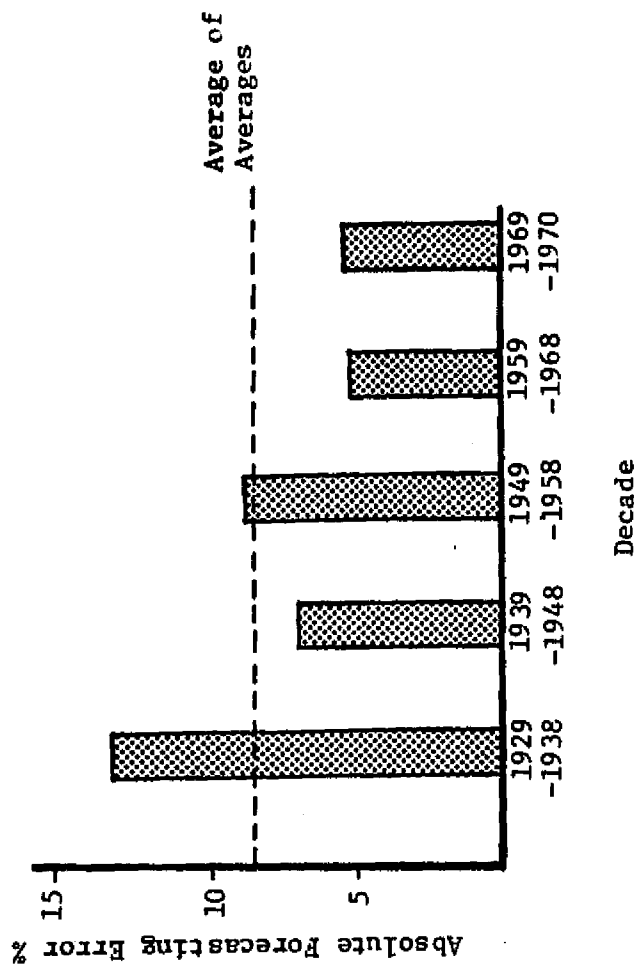FIGURE B-6

PUBLISHED USDA-SRS FORECAST ACCURACY

WINTER WHEAT – TOTAL U.S. CROP
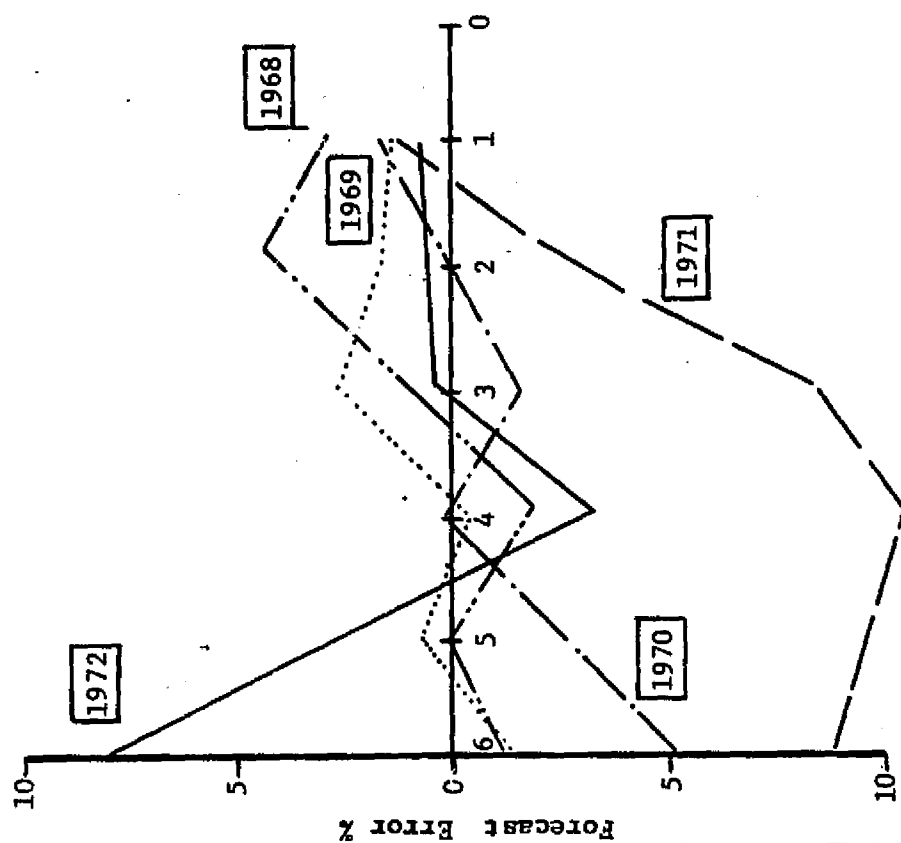
FIGURE B-8

WINTER WHEAT DATA COVER 1957-1972



FIGURE B-9

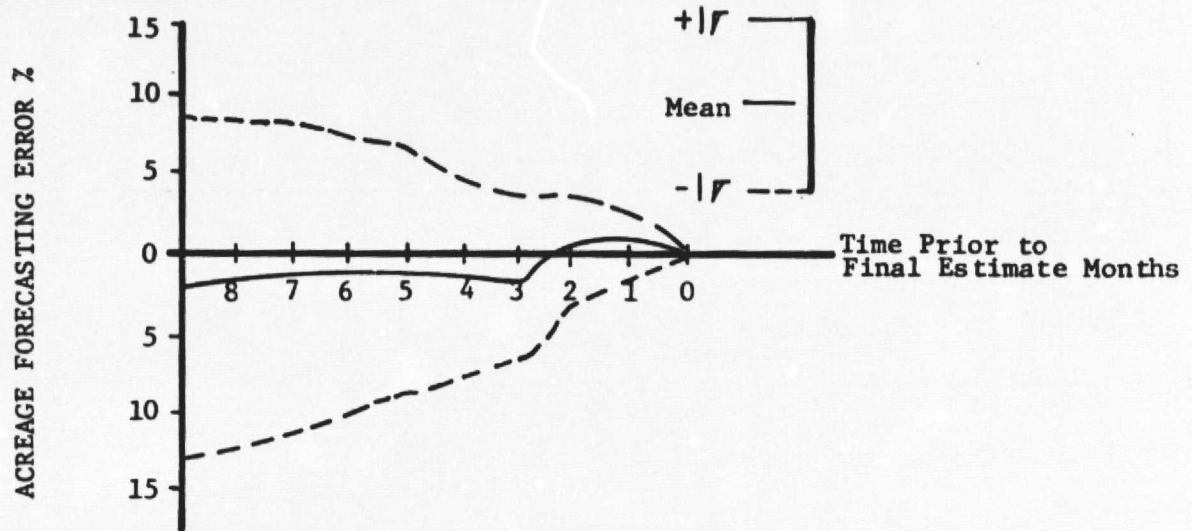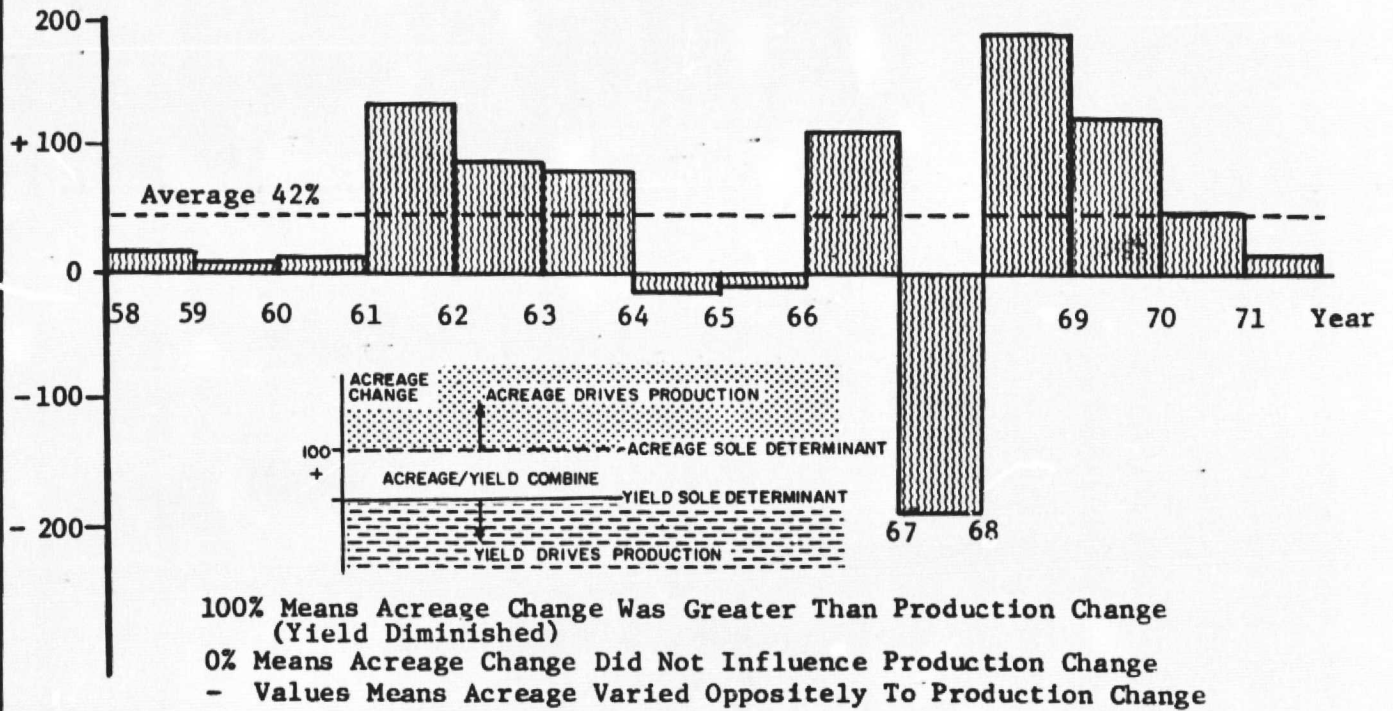ACREAGE CHANGE AS A % OF TOTAL PRODUCTION CHANGE



100% Means Acreage Change Was Greater Than Production Change
   (Yield Diminished)
0% Means Acreage Change Did Not Influence Production Change
- Values Means Acreage Varied Oppositely To Production Change

# COUNTRIES INVOLVED IN CROP SURVEY ANALYSIS

| Country | SURVEY CLASSIFICATION | | | Country | SURVEY CLASSIFICATION | | | Country | SURVEY CLASSIFICATION | | | Country | SURVEY CLASSIFICATION | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Acreage Forecast | Yield Forecast | Statistical Sampling | | Acreage Forecast | Yield Forecast | Statistical Sampling | | Acreage Forecast | Yield Forecast | Statistical Sampling | | Acreage Forecast | Yield Forecast | Statistical Sampling |
| **EUROPE** | | | | Guadalope | B | B | -- | Malaya | A | A* | S | Somaliland | D | D | -- |
| Austria | A | B | -- | Guatemala | A | A* | -- | India | A | B | -- | Togo | A | A* | |
| Belgium | A | B | -- | Martinique | B | B | -- | Indonesia | B | A | -- | Fr. West Africa | A | A* | |
| Denmark | A | A | S | Mexico | B | B | -- | Iran | D | D | -- | Gold Coast | A | A* | S* |
| Finland | A | A | S | Panama | A | A* | S | Iraq | C | C | -- | Kenya | A | A | |
| Germany FR. | A | B | S | Puerto Rico | A | A* | S | Israel | B | B | S* | Madascar | C | C | -- |
| France | A | D | -- | USA | A | A | S | Japan | A | B | S | Nigeria | D | D | S* |
| Greece | B | B | -- | **SOUTH AMERICA** | | | | Jordan | C | C | -- | Rhodesia, Northern | A | A | -- |
| Ireland | A | A | S | Argentina | A | A | S | Lebanon | C | C | -- | Nyasaland | A | A | -- |
| Italy | B | B | S | Brazil | C | C | S* | Pakistan | C | C | -- | Rhodesia, So. | A | C | -- |
| Luxembourg | A | B | -- | Chile | A | D | -- | Phillipines | B | B | -- | Morocco, Sp | A | A* | -- |
| Netherlands | A | B | S | Colombia | A | A* | B* | Syria | C | C | -- | Swaziland | A | A* | -- |
| Norway | A | B | S | Ecuador | A | A* | -- | Thailand | B | B | -- | Tanzania | D | D | S* |
| Portugal | C | C | -- | Fr. Guiana | -- | -- | -- | Turkey | B | B | S | Tunisia | A | A* | -- |
| Spain | A | B | -- | Paraguay | A | C | S* | **AFRICA** | | | | Union South A | A | A | -- |
| Sweden | A | B | S | Peru | D | D | -- | Algeria | B | B | -- | Uganda | C | C | S* |
| Switzerland | A | B | -- | Surinam | A | A* | -- | Sudan | B | B | -- | **OCEANIA** | | | |
| United Kingdom | A | C | S | Uruguay | A | A* | S | Basutoland | D | D | S | Australia | A | A | S* |
| Yugoslavia | B | B | S | Venezuela | B | B | -- | Bechuanaland | C | C | -- | Fiji | A | A | -- |
| **NORTH-CENTRAL AMERICA** | | | | **ASIA** | | | | Belgian Congo | B | B | — | Hawaii | A | A | -- |
| Canada | A | A | S | Burma | A | A | -- | Egypt | A | A* | -- | New Caledonia | A | A* | -- |
| Costa Rica | A | A* | S | Ceylon | A | B | S | Cameroons, Fr. | A | A* | -- | New Zealand | A | A | S |
| Dominican Republic | A | A | -- | Cyprus | A | A* | -- | Morocco, Fr. | C | C | -- | | | | |
| El Salvador | B | B | -- | | | | | | | | | | | | |

Explanation of Symbols:  **A.** Sampling at farm level   **B.** Sampling at commune level   **C.** Sampling at district level   **D.** Sampling at province level   **S.** Stratified Sampling
\* System under development

also tend to decrease with decreasing GNP per capita. Table I 2 synthesis the world's current crop re orting systems into three fundamental categories: Developing, Intermediate, Advanced. The agricultural data-gathering network of the US is shown in Figure B-10, that of a typical intermediate county in Figure B-11.

The basis for potential improvements in acreage measurement procedures lies in the fact that present methods, for budgetary reasons, only measure limited samples. Increased sampling, or, in the limit, wall-to-wall measurements, would improve accuracy; performance of the current sampling task without laborious field surveys would reduce costs.

As regards yield estimates, the benefits of increased sampling still apply but with significant differences. Expected yield is based to a significant degree upon the assessment of present crop condition. To the extent that this quantity can be measured more accurately by increasing the number of samples, yield forecasts will be improved. However, yield also depends upon other factors more difficult to estimate: principal among these is weather. Therefore, until the state of technology will allow accurate prediction of weather, an irreducible uncertainty will remain in the forecast. Better sampling methods can, however, materially assist in reducing the errors of estimate.

TABLE B-2

| | | | |
|---|---|---|---|
| | THE THREE BASIC LEVELS OF CROP FORECASTING SYSTEMS | | |
| | Developing | Intermediate | Advanced |
| | Typical: Dominican Republic | Typical: Italy | Typical: United States of America |
| Regional Administrative Structure | 1599 sections grouped into 69 ommunes | 7851 Communes grouped into 91 Provinces | 1700 enumerated areas grouped into 44 States |
| Crops Surveyed | Maize, rice, beans, potatoes, onions, garlic, peanuts, coconuts, oranges, bananas, cocoa, sugar cane, plantains, pineapple, coffee (in pod), avocado, pears, cotton, tobacco | Wheat, rye, barley, oats, maize, rice, sugar beets, potatoes, peas, beans, vineyards, fruits, olives, linseed rapeseed, vegetables, tobacco, fibers | Grains, fodder crops, tuber and root crops, sugar crops, pulses, oilseeds, hay and grass seeds, vegetable seeds, fruits, nuts, vegetables, tobacco, fibers |
| Methods of Data Gathering | • Interview of producers by enumerators | • Crop area from personal judgement, supplemented by cadastral survey<br>• Crop yields from local inquiry | • Direct inquiry to farmer respondents<br>• Enumerators with aerial photos<br>• Objective yield measurement by specialists |
| Sampling Structure | None | Simple sampling procedure | Multi-frame stratified sampling procedures |
| Organization | Municipal Statistical Board in each commune | • Data from commune collected by local correspondent assisted by provincial agricultural inspector<br>• Central Institute of Statistics issues technical directives and publishes results | • SRS HQ staff supported by 44 State offices comprising 9 crop regions.<br>• Refer to Figure 9 |
| Frequency of Crop Reporting | Every 3 months | Two crop reports per year First estimate at planting time; second estimate at harvest | Multiple crop reports per year. Intentions to plant-yearly per crop. Acreage, crop condition, production forecast - monthly for 3-6 months. Final production and yield-yearly. |

FIGURE B-10

U.S. GLOBAL CROP FORECAST DATA & INFORMATION NETWORK - 1977
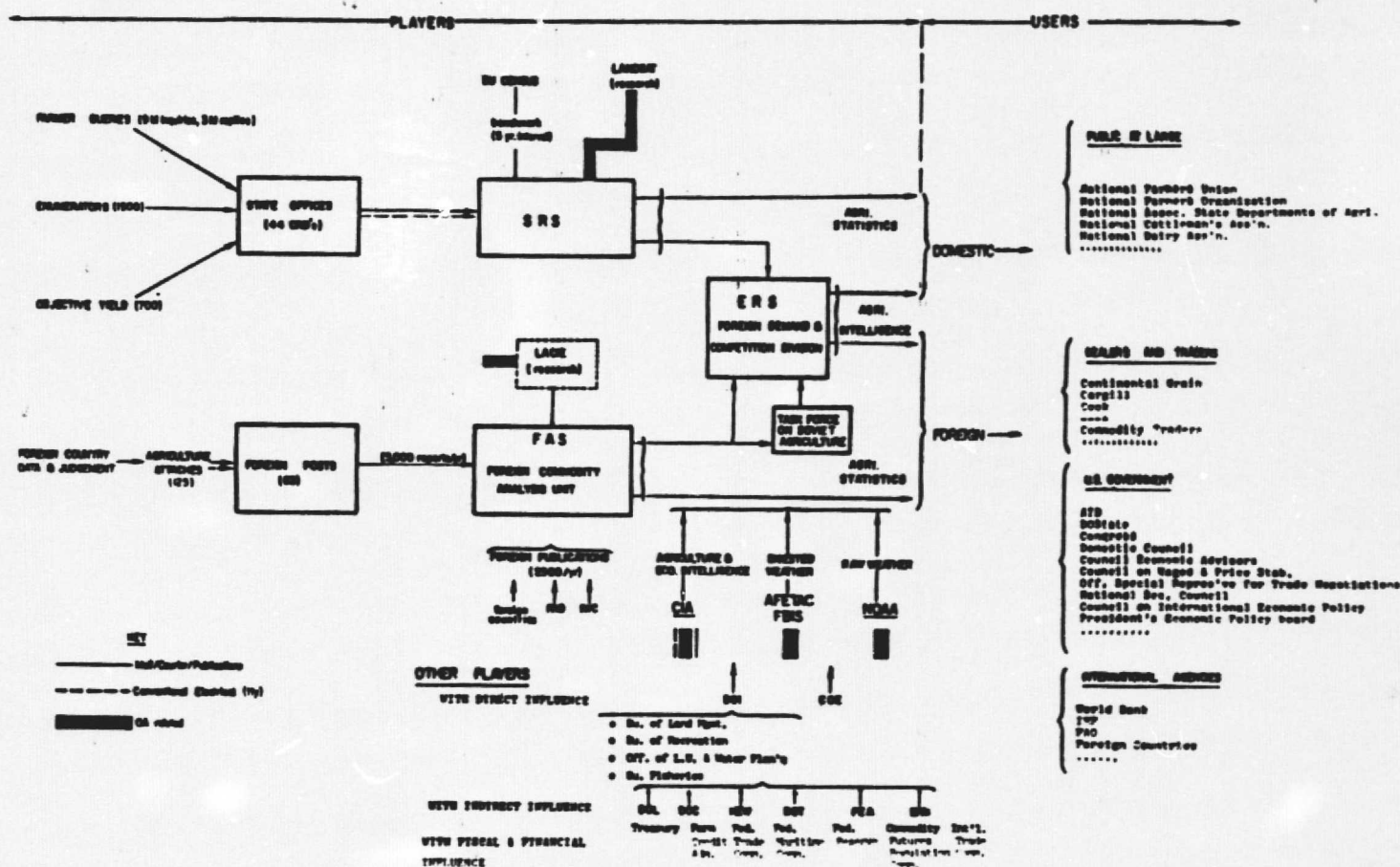
## FIGURE B-11
## TYPICAL FOREIGN COUNTRY CROP INVENTORY & FORECAST INFORMATION NETWORK (ITALY)

Acreage:
Statistical Sampling
(5,300 segments,
stratified, 750
enumerators)

Teletype

Domestic
Information

**Public-at-large**
National Farmers Association
National Cattle Growers Co-op
Chambers of Commerce

**Dealers and Traders**
AIMA (Semi-Gov't. Purchasing
Agency)
ISVET (Private Forecasting
Group)
Private Traders

| Provincial Inspectorates of Agriculture (92) | Regional Inspectorates (18) | Central Statistical Institute |
|---|---|---|

**Government Agencies**
Ministry Agriculture &
Forestry

. . . . . .

**Foreign Agencies**
USDA Attache'
Other Attache's

Yield:  Subjective
plus Farmer
Queries (same per-
sonnel)

National
Weather
Service

**International Agencies**
FAO

— Mail/Courier/Publications

-- Conventional electrical

# APPENDIX C

# DETERMINATION OF THE ACCURACIES OF AGRICULTURAL ACREAGE

# ESTIMATION ACHIEVED BY ORBITING DATA SYSTEMS

# (LANDSAT)

## APPENDIX C

## DETERMINATION OF THE ACCURACIES OF AGRICULTURAL ACREAGE ESTIMATION

## ACHIEVED BY ORBITING DATA SYSTEMS

### (LANDSAT)

### 1.0  Accuracies Achieved by ERTS Investigators

The ERTS crop classifications represent an intensive series performed during the early phases of the LANDSAT program from its launch in July 1972 until 1974.  The spacecraft was at that time named ERTS, later changed to LANDSAT.

The available principal investigations reporting numerical results were used in the analysis [16] [17] [18] [19] [20] [21] [22] [23] [24].

Table C-1 synopsizes the investigations analyzed.  Nine ERTS investigators performed 224 distinct tests in ten locations, for seven major crop types.  Thirty-two of these experiments included sufficiently complete description of the corresponding ground truth for complete comparison.  Forty--three provided aggregate results against various estimates of the ground truth e.g. SRS estimates, State Yearbook data etc.  The balance indicated only the percent of pixels correctly identified.

The cumulative curve of the results, presented in Figure C-1A, graphs the accuracies achieved versus the percentage of the tests achieving them. It presents the results for both the inventory and the mapping, or land use,

TABLE C-1

SUMMARY OF ERTS-1 ACREAGE MENSURATION RESULTS

| INVESTIGATOR | PLACE | DATE | MAJOR CROPS | AVERAGE FIELD SIZE | METHOD OF COMPUTATION | RANGE OF ERROR |
|---|---|---|---|---|---|---|
| R. Bryan Erb, NASA (87) | Hill Co., Montana | 8/7/72 | Wheat | 42.8 Ha. | | |
| | Holt Co., Nebraska | 7/29/72 | Corn | 366.3 Ha. | Planimeter | 0.1-35% |
| | Imperial Co., Cal. | 11/6/72 | Asparagus | 45.9 Ha. | | |
| | Butte Co., Cal. | 9/19/72 | Rice | 255.2 Ha. | | |
| D.S. Simonett | Williams Co., N.D. | 6/5/73 | Spring Wheat Durum Oats Barley | 15.8 Ha. | Pixel Count (Proportion to Maximum) | 0-25% |
| Craig L. Wiegand, ARS (31) | Hidalgo Co., Tex. | 1/21/73 5/27/73 | Vegetables Citrus Cotton Sorghum Rangeland | No-64.8 Ha. So- 8.9 Ha. Cen-6.9 Ha. Tot-12.2 Ha. | Pixel Count | 2-350% |
| Marion F. Baumgardner, LARS (5) | Lubbock Co., Tex. Greeley Co., Kan. Humboldt Co., Iowa | 6/18/73 7/24/73 6/19/73 8/26/73 | Wheat Cotton Sorghum Wheat Pasture Corn Soybeans | NA | Pixel Count (Proportion to Maximum) | 1-35% |
| David A. Landgrebe, LARS (15) | 3 Counties in Ill. | 8/9/72 | Corn Soybeans | NA | Bias Corrected & Uncorrected ERTS results determined on basis of equal & non-equal weights. | 11-65% |
| | 12 Counties in Ind. | 8/21/73 | Corn Soybeans | NA | Pixel Count | |
| Humberto C. Garuti, Argentina (6) | Pampus of Argentina | 10/5/72 | Wheat Barley Rye Corn | NA | Pixel Count | 1-55% |
| J.P. Mahlstede, Iowa State (10) | Ames, Iowa Doon, Iowa | 8/72 | Corn Alfalfa Soybeans | NA | Planimeter | 7-40% |
| Ralph Shay, NASA (14) | Indiana Illinois | -- | Corn Soybeans | 6-17 Ha. | Pixel Count (Pro. to Max) | 0.5-21% |
| Gene A. Thorley, Univ. of Calif. (22) | San Joaquin Co., Calif. | 7/24/72 | Range Vegetables Fruit | 17 Ha. | Pixel Count (Proportion to Maximum) | 0-380% |

Numbers in parentheses indicate the number of tests performed by each investigator.
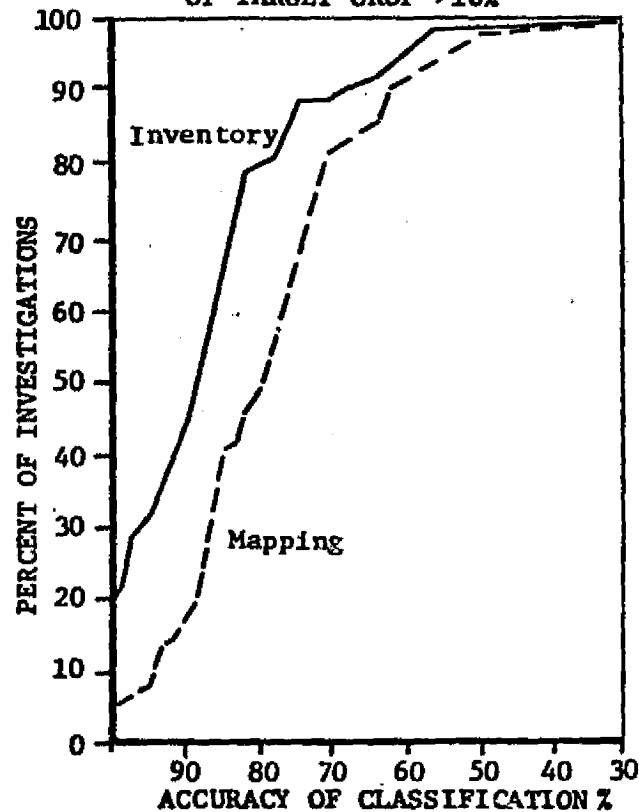
# FIGURE C-1

## DISTRIBUTION OF CLASSIFICATION ACCURACIES FOR ERTS INVESTIGATIONS ON ACREAGE MENSURATION

### A. CLASSIFICATION ACCURACY OF ERTS INVESTIGATIONS
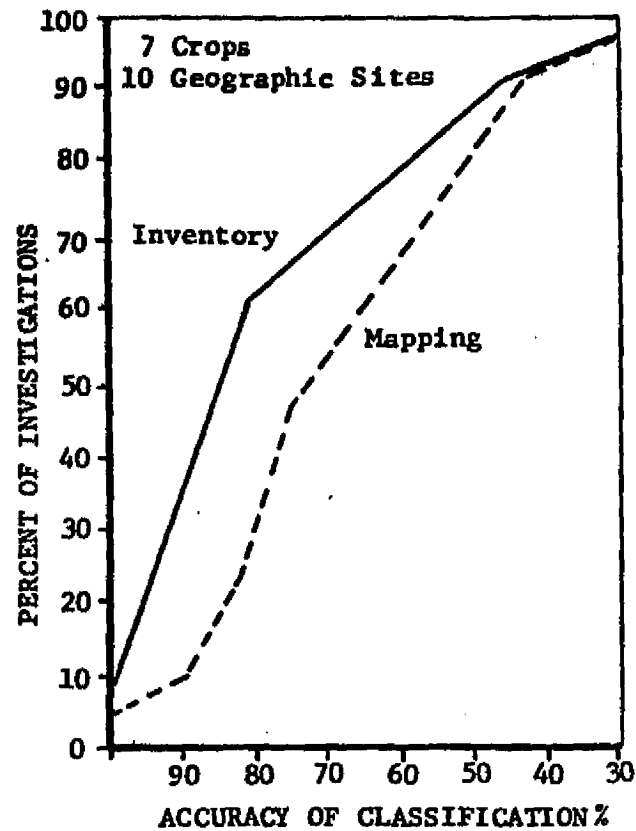


NUMBER OF DATA
POINTS: 224

### B. CLASSIFICATION ACCURACY OF ERTS INVESTIGATIONS WITH PROPORTION OF TARGET CROP >10%



NUMBER OF DATA
POINTS: 64

mode. Inventory mode means the determination of the proportion of a given crop within the area under examination, without however specifying the crop's geographic location. It is the mode universally used for estimating agricultural acreage. Mapping mode, as the word indicates, connotes the identification not only of the species sought, but also of their geometric boundaries and geographic locations. This mode is generally not employed in conventional agricultural acreage estimation, due its higher cost with respect to the inventory mode.

Note that the accuracies achieved in the mapping mode are lower than those competing to the inventory mode: this is quantitatively consistent with the fact that the mapping mode requires more information than proportion estimation: conversely, for equal informatiion content, its level of performance should be lower.

The accuracy of multispectral classification worsens when the proportion of target crops to the total scene is small: a subset was extracted from the total statistical population, based on the criterion of 10% or larger proportion of target crops. This yielded 64 distinct data points: the test results are presented in Figure C-1B. Note the higher accuracies achieved.

A breakdown of the mode – whether inventory or land use – employed for each experiment by each ERTS investigator is given in Table C-2.

TABLE C-2

BREAKDOWN OF THE NO. OF EXPERIMENT VARIANTS FOR EACH ERTS
INVESTIGATOR BY MODE

| INVESTIGATOR | LAND USE | INVENTORY | LAND USE PLUS INVENTORY | TOTAL |
|---|---|---|---|---|
| LANDGREBE | 51 | 2C | 6 | 86 |
| ERB | 16 | 0 | 13 | 29 |
| GARUTI | 1 | 8 | 0 | 9 |
| WIEGAND | 4 | 6 | 0 | 10 |
| BAUER | 76 | 0 | 0 | 76 |
| EARTHSAT | 0 | 0 | 1 | 1 |
| MAHLSTEDE | 0 | 0 | 2 | 2 |
| BAUMGARDNER | 1 | 0 | 0 | 1 |
| THORLEY | 0 | 0 | 10 | 10 |
| TOTAL | 149 | 43 | 32 | 224 |

One-hundred sixty-nine experiments reported the individual field areas. The distribution of the field sizes shown in Figure C-2 indicates the predominance of fields below 50 hectares and above 10,000 hectares.

The sensitivity of the classification to area is indicated in Figure C-3 for each of the area classes given in Figure C-2. The general trend is that the mean error within each class is inversely related to the area. Note however that there is a wide dispersion of errors for each area class: this is indicative of the wide dispersion in the fraction of the total scene comprised by the crops investigated. Figure C-4 summarizes the results for the 169 investigations.

FIGURE C-4

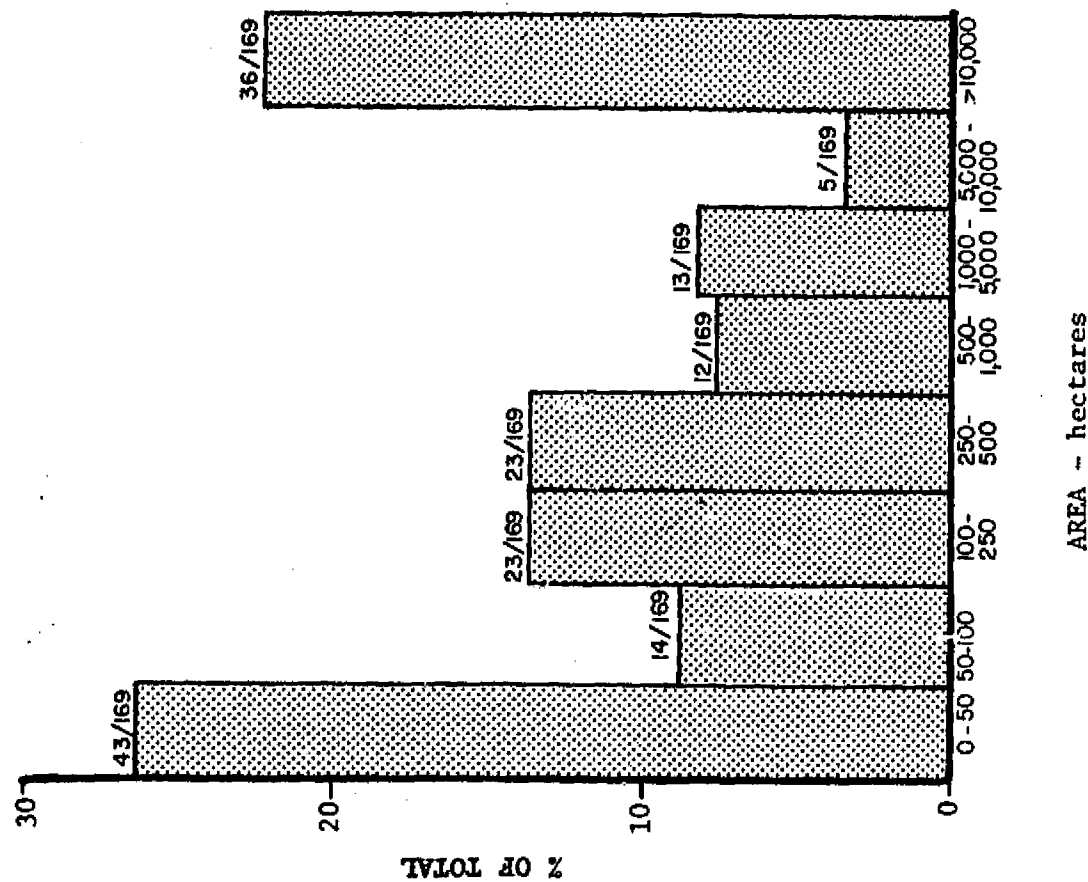MENSURATION FROM ERTS – DISTRIBUTION OF ERRORS, INVENTORY MODE
169 DATA POINTS



FIGURE C-2

MENSURATION FROM ERTS – DISTRIBUTION OF AREAS MENSURATED
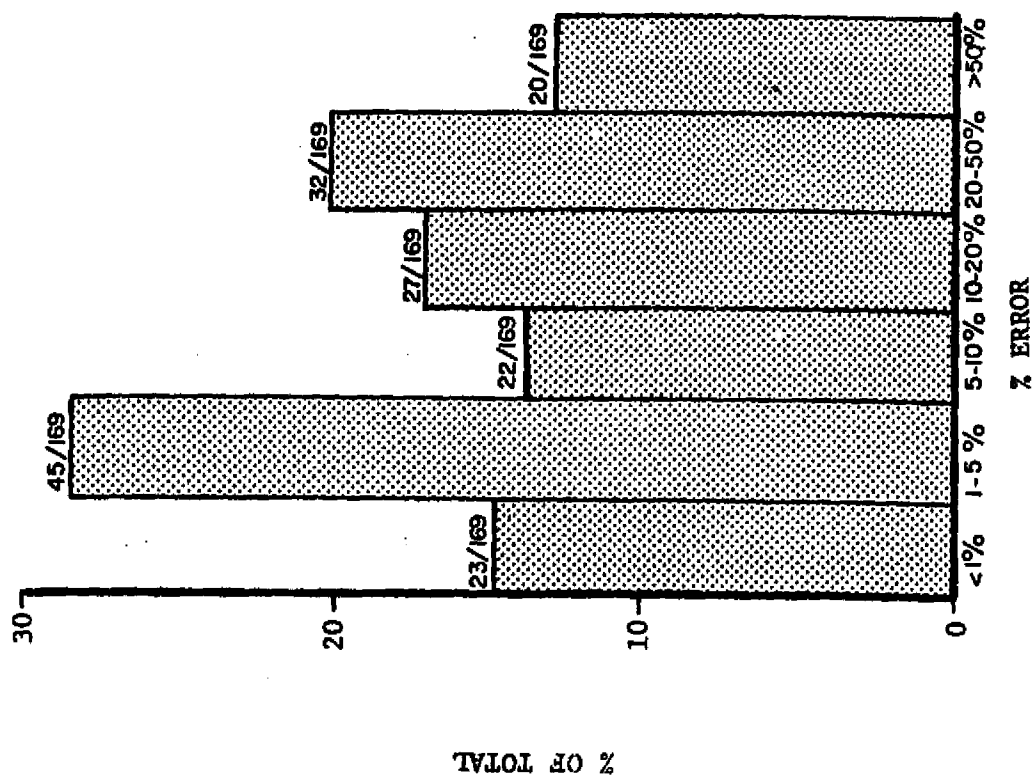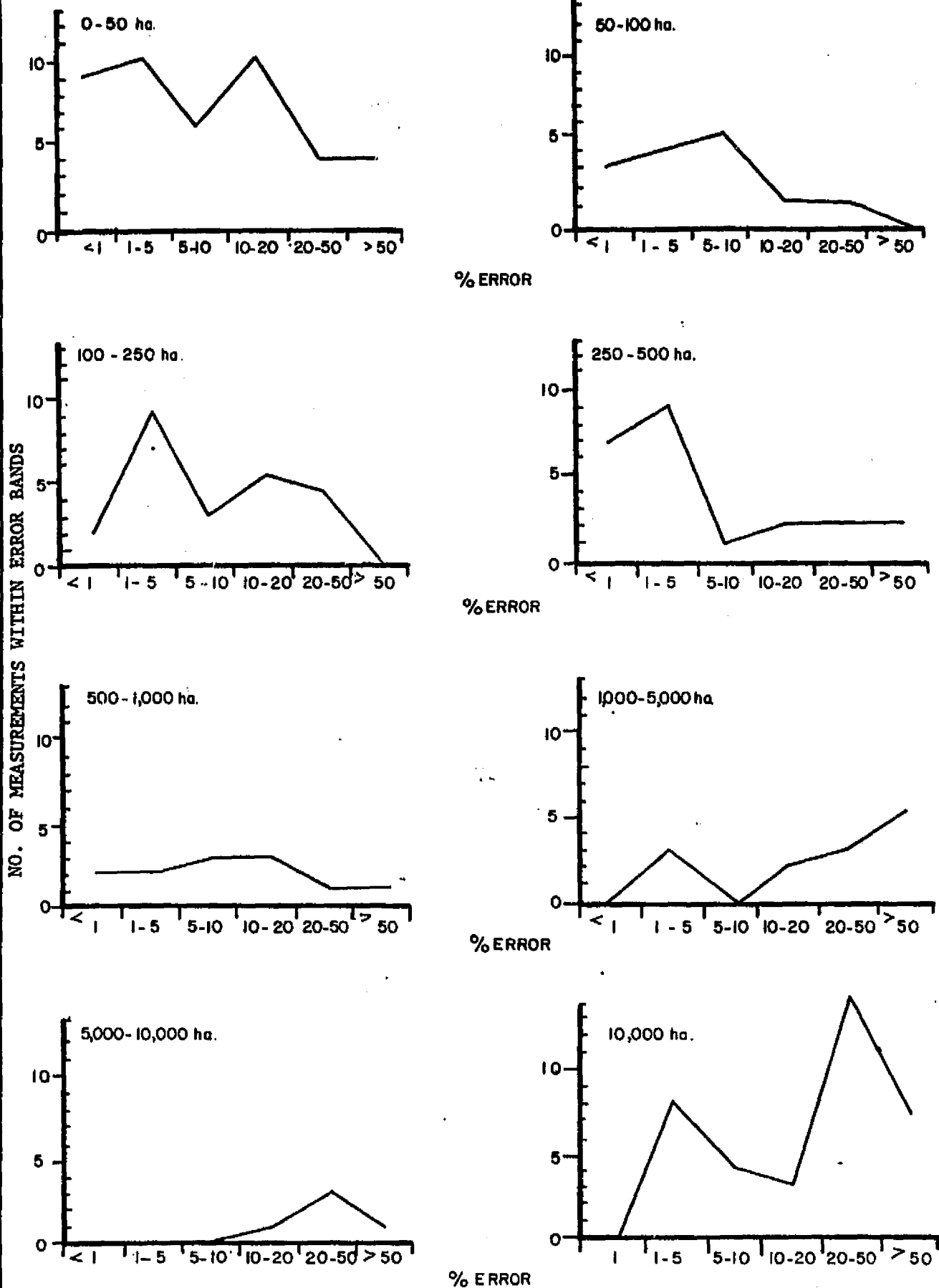169 DATA POINTS

# FIGURE C-3
## SENSITIVITY OF EPTS CLASSIFICATION TO AREA — INVENTORY MODE
### 169 DATA POINTS

The effect of small crop acreages was eliminated by reducing the set of investigations analyzed from 169 to 64 as indicated in Table C-3. Figures C-5 and C-6 present the distributions of investigations as a function of percent error for the 64 investigations with acreage fraction of crop greater than 10%.

```
+-----------------------------------------------------------------------+
|                             TABLE C-3                                 |
|                                                                       |
|     AGRICULTURAL CROP CLASSIFICATION FROM ERTS INVESTIGATIONS REDUCED TO |
|                  PROVIDE ACREAGE FRACTION > 10%                       |
|                                                                       |
|                                                                       |
|   NUMBER OF INVESTIGATIONS SELECTED FOR THIS SAMPLE:  32 WITH 64 CROP DATA |
|                                                       POINTS          |
|                                                                       |
|   NUMBER OF INVESTIGATORS PERFORMING THE 32 INVESTIGATIONS ANALYZED:  5 |
|                                                                       |
|   RANGE OF AREAS CLASSIFIED:  400 TO 100,000 HA:  MEDIAN 1700         |
|                                                                       |
|   GEOGRAPHIC LOCATIONS:  9 STATES:  (IL, IN, MT, IA, CA, NE, GA, TX, ND): |
|                          APPROXIMATELY 20 COUNTIES                     |
|                                                                       |
|   PRINCIPAL CROPS SELECTED:  WHEAT, SMALL GRAINS, CORN, SOYBEANS, BARLEY, |
|                              PASTURE, ALFALFA                          |
|                                                                       |
|   AVERAGE NUMBER OF DISTINCT CROPS CLASSIFIED PER INVESTIGATION:  3 TO 4 |
|                                                                       |
|   NUMBER OF DISTINCT CROPS SELECTED FOR COMPARISON:  AVERAGE 2 PER INVES- |
|   TIGATION.  CRITERION:  ONLY CROPS WHOSE ACREAGE RATIO TO TOTAL CROP  |
|   ACREAGE EXCEEDED 10%                                                 |
|                                                                       |
|   THESE YIELDED 64 DATA POINTS ON PRINCIPAL CROPS.                     |
+-----------------------------------------------------------------------+
```

The classification results are presented in Table C-4 for both the 224 and 64 investigation data sets. It shows that the one sigma accuracy of classification in the *inventory mode* for the total population of 224 data points is 74%: this increases to 86% for the subset of 64 data points featuring 10% or greater proportions.

# DISTRIBUTION CLASSIFICATION ERRORS

## TOTAL NUMBER OF INVESTIGATIONS = 64
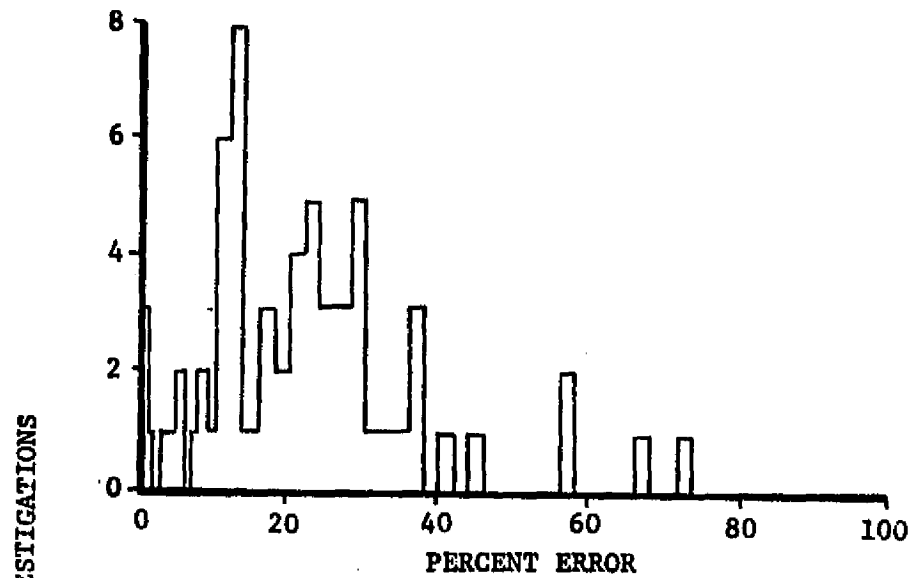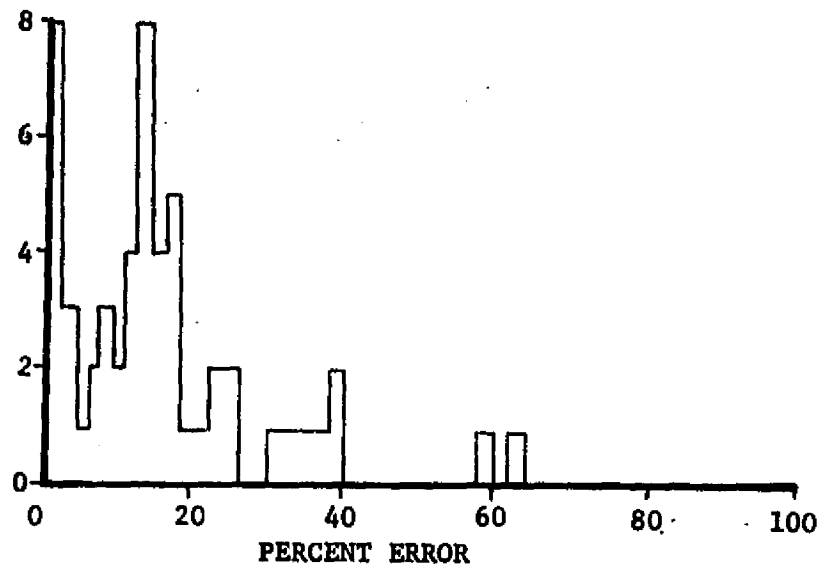### ALL CROPS >10% OF SCENE

FIGURE C-5

LAND USE MODE



FIGURE C-6
INVENTORY MODE

**TABLE C-4**

**ACCURACY OF CLASSIFICATION OF CROPS ACHIEVED**

**BY ERTS INVESTIGATORS**

UNIVERSE OF 224 CLASSIFICATION TESTS ON 7 CROPS, 10 GEOGRAPHIC LOCATIONS, ONE SIGMA CONFIDENCE

|  | PROPORTION ESTIMATION | MAPPING |
|---|---|---|
| ACCURACY % | 74 | 63 |
| ERROR % | 26 | 37 |

64 CLASSIFICATION TESTS EXTRACTED FROM THE ABOVE UNIVERSE, WHEREIN THE CROP-TO-TOTAL-AREA PROPORTION WAS GREATER THAN 10%.

| | | |
|---|---|---|
| ACCURACY % | 86 | 76 |
| ERROR % | 14 | 24 |

## 2.0 Accuracies Achieved in Acreage Estimation after 1974

ERTS investigations ceased after 1974. To assess whether the early performance has improved in the approximately four years elapsed since, a number of subsequent investigations, performed by reputable organizations, were analyzed. Table C-5 is a summary of the key Landsat investigations dealing with crop classifications reported since 1974.

Figure C-7 presents the results of the investigations in the inventory and land use mode. The one sigma accuracy was 86.3% for the inventory mode and 73.4% for the land use mode respectively. The overall results are also shown in Table C-6.

## TABLE C-5

### SUMMARY OF KEY LANDSAT CLASSIFICATION INVESTIGATIONS SINCE 1974

| DESCRIPTION | LOCATION | NUMBER OF CLASSES | AVERAGE INVENTORY ERROR | AVERAGE LAND USE ERROR | DATE | PROCEDURE |
|---|---|---|---|---|---|---|
| Area estimation of crops by digital analysis of Landsat Data, LARS, Purdue University | Kansas 4,613,000 Ha. | 2 (wheat, other) | 1.5% compared with SRS estimate | - | 1977 | Computer classification of wheat. |
| High-altitude vs. Landsat imagery for digital crop identification (LACIE) | Kern County, California 11,839 acres | 10 (cotton, grapes, melons, tomatoes, sugarbeets, wheat, fallow, oranges, natural vegetation, other) | 24% | 22% | 1977 | Computer crop classification |
| Pattern Recognition of soils & crops from space, USDA | Imperial Valley, California (53 fields) | 5 (alfalfa, barley, sugarbeets, bare soil, salt flats) | - | 28% | 1975 | Computer crop classification |
| Land use inventory for Alto Lazio, ERIM | Alto Lazio, Italy (53,284 ha.) | ·5 (Piscolo, Seminativi, Foreste, Orchards, Urban) | 12.3% | - | 1977 | Manual crop class from Landsat images |
| Landsat spectral signatures: studies with soil associations and vegetation. South Dakota Agricultural experiment station | Brookings Co. South Dakota 12,950 ha. | 4 (corn, small grain, grass, water) | 6.1% | 6.3% | 1978 | Computer classification of Landsat imagery |
| Cost, accuracy, and consistency comparisons of land use maps made from high-altitude aircraft photography and ERTS imagery. | Central Atlantic | 6 (urban, water, agricultural, forest, non-forested wetland | 12.1% | 30.5% | 1975 | Manual classification from Landsat imagery |
| Landsat urban area delineations. Intralab Project 75-3 | Prince Georges Co., MD 84,100 ha. | 2 (residential, other) | 27% | 27% | 1977 | Computer classification |
| Global agricultural productivity evaluation from Landsat data, J. Schubert, Gregory Geoscience Ltd. | 4 Canadian test sites and acres analyzed 1) Swiftcurrent (1844 acres) | 2 (wheat, other) | 0% | - | 1976  July 1973 | Manual classification from enhanced Landsat imagery |
| | (7204 acres) | | 0.6% | - | July 1975 | |
| | (6960 acres) | | 3.4% | - | July 1976 | |
| | 2) Raymond (2175 acres) | 2 (wheat, other) | 3.8% | - | July 1974 | |
| | (4811 acres) | | -4.7% | - | July 1976 | |
| | 3) Melford (4206 acres) | " | -13.1% | - | Aug. 1976 | |

TABLE C-5  (cont'd)

SUMMARY OF KEY LANDSAT CLASSIFICATION INVESTIGATIONS SINCE 1974

| DESCRIPTION | LOCATION | NUMBER OF CLASSES | AVERAGE INVENTORY ERROR | AVERAGE LAND USE ERROR | DATE | PROCEDURE |
|---|---|---|---|---|---|---|
| | 4) Stoney Mountain (3439 acres) | 2 (wheat, other) | 7.2% | - | Aug. 1974 | |
| Crop identification technology assumed for remote sensing (CITARS), B.J. Davis, A.H. Feiueson, NASA, JSC, Houston, Texas | Texas | 3 (corn, soybeans, other) | 13.2% | - | 1975 | Computer classification |
| Land use classification for 6 Rocky Mountain states - using Landsat multispectral - multi-temporal data | Fox Creek, Colorado (1268 acres) | 11 (Rabbitbrush, grass/rabbitbrush, dense shrub, mahogany, meadow, wet pasture, doug fir, Pinyon-Juniper, Ponderosa Pine, Aspen, cottonwood) | - | 25.6% | 1977 | Computer classification |
| | Fox Creek, Colorado (1268 acres) | 5 (shrub land, non-irrigated grassland, irrigated grassland, coniferous forest, deciduous forest) | - | 17.8% | | |
| | | 2 (Rangeland, forest) | - | 10.5% | | |

## FIGURE C-7

LANDSAT POST 74 CROP CLASSIFICATION RESULTS

TOTAL NO. IN SAMPLE = 18
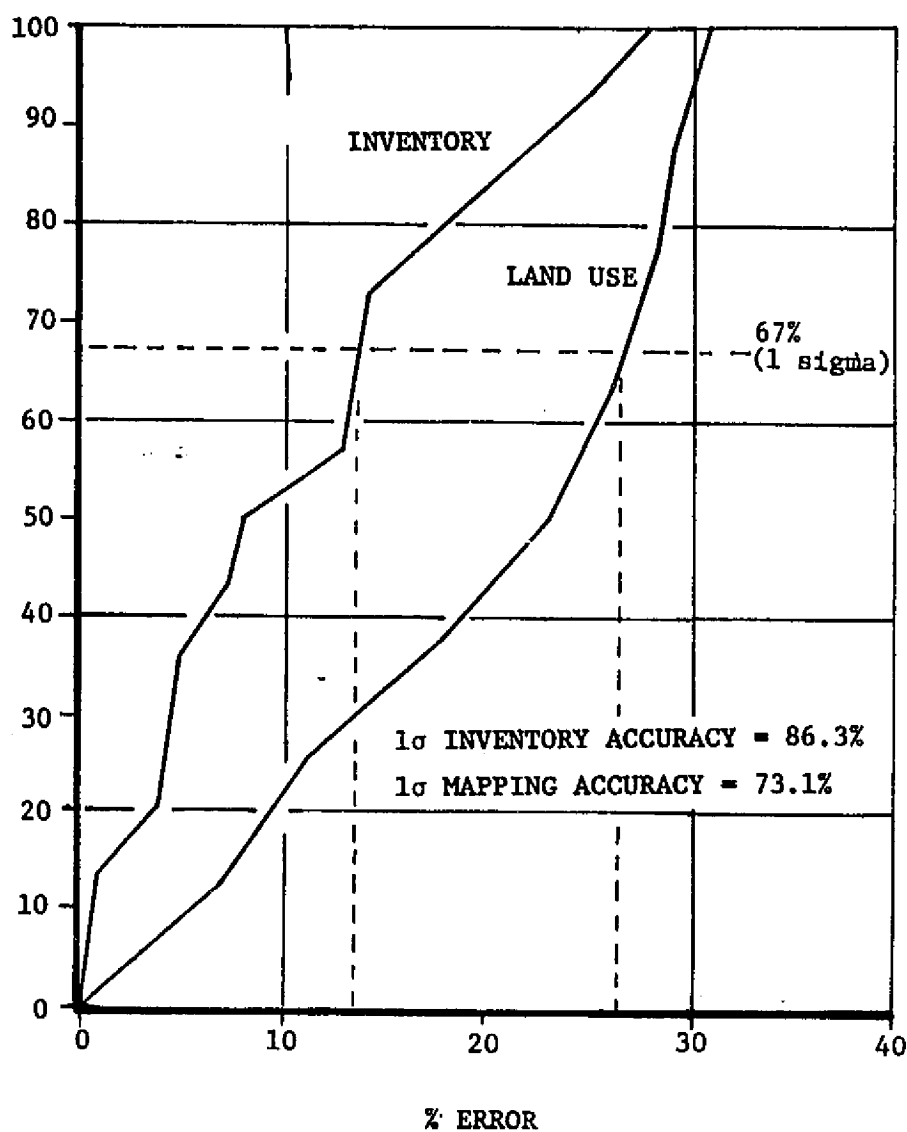


% ERROR

1 SIGMA INVENTORY ERROR ⁓13.7%
1 SIGMA LAND USE ERROR ⁓26.6%
MEDIAN ABSOLUTE INVENTORY ERROR ⁓ 8%
MEDIAN ABSOLUTE LAND USE ERROR ⁓23%

---

**TABLE C-6**

**ACCURACY OF CLASSIFICATION OF CROPS ACHIEVED FROM 1974 TO 1978**

Universe of 18 classification endeavors on 10 crops, 13 geographic locations, one sigma confidence.

|  | PROPORTION ESTIMATION | MAPPING |
|---|---|---|
| Accuracy % | 87% | 73% |
| Error % | 14% | 27% |

---

## 3.0 Description of the Experiments in ERTS Crop Classification

Tables C-7 through C-14 present the highlights of each ERTS experiment: the crops analyzed, their locations, the dates of the ERTS observations, the method of analysis employed.

Figures C-8 through C-16 depict the scattergrams of the differences between measured values and ground truth. The scattergrams present for each group of investigations, the error experienced in measuring each crop area versus the actual areas. The solid downward sloping straight lines represent the error of mensuration $e_m$% expected from the ERTS pixel size. on the assumption of perfect discrimination. They obey the relationship:

$$e_m\% = \frac{2\,r}{\sqrt{A}}$$

Where:

r is the average dimension of the ERTS pixel 70 meters

A is the area mensurated in hectares

## DETAILED SUMMARY OF THE ERTS CROP MENSURATION CASES

## TABLE C-7

### HIGHLIGHTS OF ERTS EXPERIMENT

Source:  The ERTS-1 Investigation (ER-600):  Volume I - ERTS-1 Agricultural Analysis, R. Bryan Erb, NASA, August 1974.

Scene Content:

| Hill County | Summer Fallow & Sod | 55% | Imperial County | Bare Soil | 40% |
|---|---|---|---|---|---|
| | Spring Wheat | 24% | | Asparagus | 34% |
| | Winter Wheat | 19% | | Alfalfa | 15% |
| | Barley | 2% | | Carrots | 5% |
| | | | | Melons | 4% |
| | | | | Light Veg. | 2% |

| Holt County | Field Corn | 67% | Butte County | Rice | 52% |
|---|---|---|---|---|---|
| | Grass & Pasture | 16% | | Fallow | 41% |
| | Sunflowers | 4% | | Corn | 6% |
| | Popcorn & Field Corn | 4% | | Alfalfa | 1% |
| | Popcorn | 4% | | Sugar beets | .5% |
| | Field Corn & Grain Sorghum | 3% | | | |
| | Disked Fallow | 2% | | | |

Time of Measurement:

Hill County 8/7/72
    Spring Wheat - Ripening
    Winter Wheat - Harvesting
    Barley - Harvesting

Imperial County 11/6/72
    Alfalfa - Growing

Holt County 7/29/72
    Field Corn - Tasseling
    Grass & Pasture (Native Hay) - Harvesting
    Sunflowers - Heading
    Grain Sorghum - Heading

Butte County 9/19/72
    Rice - Maturing & Harvesting
    Corn - 100% Ground Cover
    Sugar beets - Harvesting

Average Field Size:

Hill County - 42.8 Ha.

Imperial County - 45.9 Ha.

Holt County - 366.3 Ha.

Butte County - 225.2 Ha.

Method of Computation:

Mechanical/electronic planimeter to manually trace field borders.

Additive color-enhanced ERTS-1 imagery of about 1:130,000 scale used as base photography for measurements.

Universal Transverse Mercator coordinate system with grid prepared for enhanced ERTS-1 imagery of about 1:130,000 scale to locate fields.

## FIGURE C-8

### EXPERIMENTAL ERRORS IN GEOMETRIC MENSURATION FROM LANDSAT IMAGERY - ERB



GROUND TRUTH MEASUREMENT - hectares
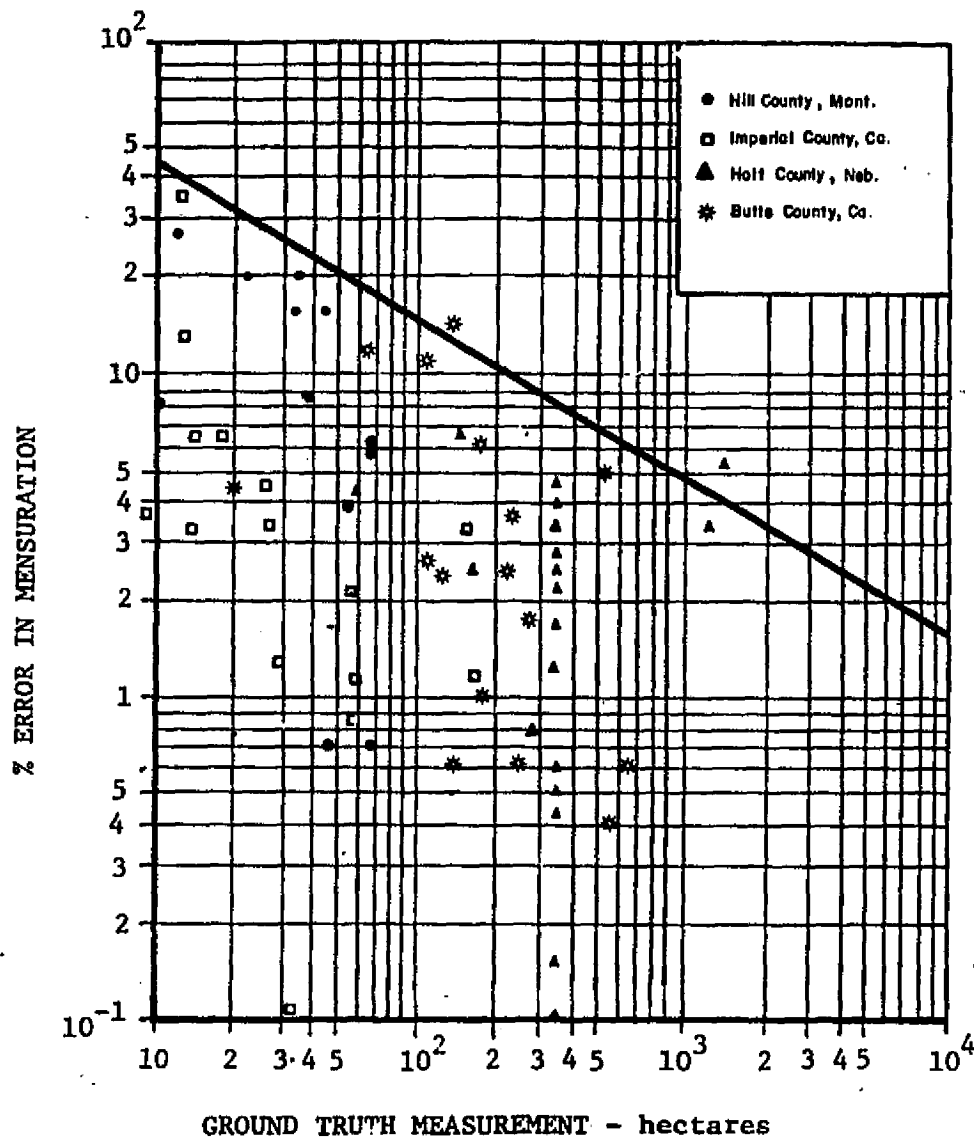
## TABLE C-8

## HIGHLIGHTS OF ERTS EXPERIMENT

**Source:** **Reflectance of Vegetation, Soil, and Water**, Craig L. Wiegand et al.,
Agricultural Research Service, USDA, November 1974.

**Scene Content:**
Urban
Agricultural
   Vegetables
   Citrus
   Cotton & Sorghum
   Idle Cropland
   Dry Debris
Rangeland
   Grass
   Shrub
   Non-agricultural
Water

**Time and Place of Measurement:**

January 21 and May 27, 1973, Hidalgo County, Texas.

**Average Field Size:**

Total County - 12.2 Ha.
Northern Region Only (Range and Pasture) - 64.8 Ha.
Central Region Only (General Farming & Citrus Growing) - 6.9 Ha.
Southern Region Only (Winter Vegetables) - 8.9 Ha.

**Method of Computation:**

Ground truth information was obtained on several items pertaining to crops,
crop histories, and farming practices by interviewing approximately 400
farmers. Typical ground truth included crop species, stage of maturity,
row spacing, plant height, percent ground cover, row direction, Munsell
color of soil, recent cultural practices, and other qualitative information.

A classification was determined for every pixel in Hidalgo County (849,000
pixels in January and 948,000 pixels in May), and a comparison was made,
using Student's t-test, between actual and computer acreage estimates of
Hidalgo County for the categories of vegetable, citrus, cotton and sorghum,
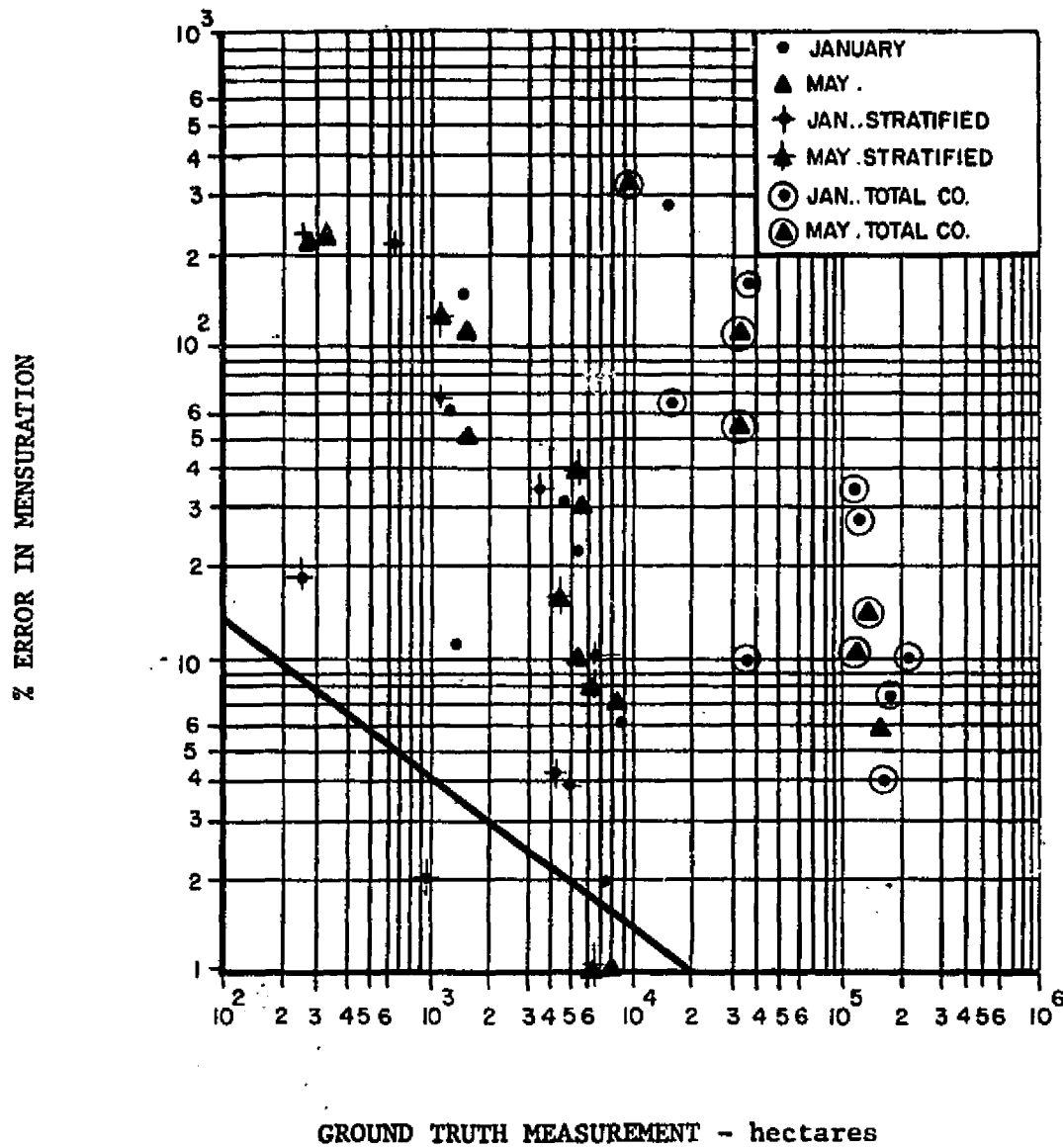idle cropland, grass, and mixed shrub.

## FIGURE C-9

### EXPERIMENTAL ERRORS IN GEOMETRIC MENSURATION FROM LANDSAT IMAGERY - WIEGAND



GROUND TRUTH MEASUREMENT - hectares

## TABLE C-9

## HIGHLIGHTS OF ERTS EXPERIMENT

Source: Evaluation and Comparison of ERTS Measurements of Major Crops and Soil Associations for Selected Test Sites in the Central United States, Marion F. Baumgardner and Staff, LARS, Purdue University, March 1974.

Scene Content:

Lubbock Co., Texas   -  Wheat - harvesting
                                 Cotton
                                 Grain Sorghum
                                 Water

Greeley Co., Kan.   -  Hard Red Winter Wheat - heading
                                 Fallow (Bare Soil)
                                 Permanent Pasture

Humboldt Co., Iowa   -  Corn
                                 Soybeans
                                 Other Row Crops

Time of Measurement:

Lubbock Co., Texas - June 18 and July 24, 1973
Greeley Co., Kan.  - June 19, 1973
Humboldt Co., Iowa - August 26, 1973

Method of Computation:

Lubbock Co., Texas - Acreage estimates were made from percentage of points in each class. The total area in the county was multiplied by these percentages to obtain an estimate of area in each cover type.

Greeley Co., Kan.  - Area of wheat was calculated from classification results.

Humboldt Co., Iowa - Area of crops was calculated from classification results.

## FIGURE C-10

### GEOGRAPHICAL LOCATION OF TEST SITES IN THE CENTRAL UNITED STATES – BAUMGARDNER
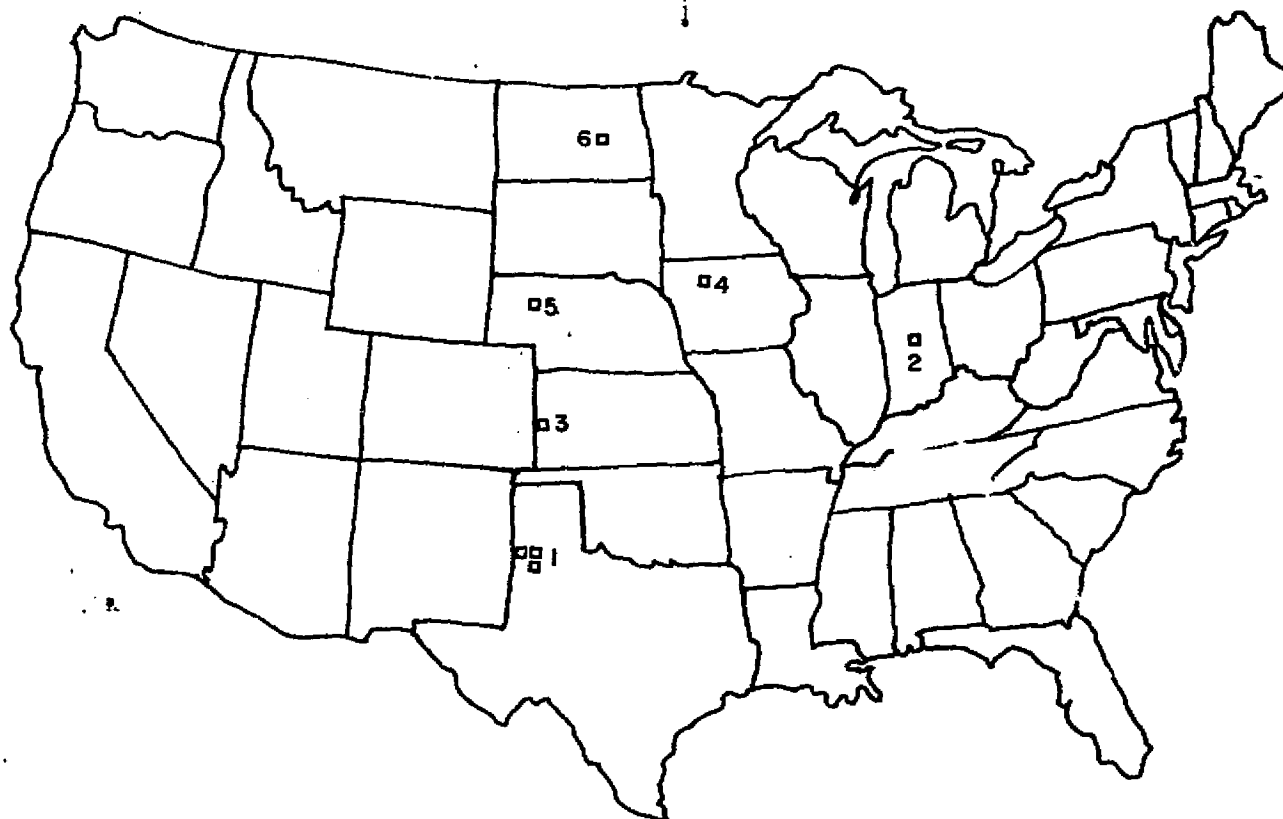


| | |
|---|---|
| 1  LUBBOCK, TEXAS REGIONAL TEST SITE | 4  HUMBOLDT COUNTY, IOWA |
| 2  BOONE AND HENDRICKS COUNTIES, INDIANA | 5  MCPHERSON COUNTY, NEBRASKA |
| 3  GREELEY COUNTY, KANSAS | 6  WELLS COUNTY, NORTH DAKOTA |

# FIGURE C-11

EXPERIMENTAL ERRORS IN GEOMETRIC MENSURATION FROM LANDSAT IMAGERY - BAUMGARDNER



GROUND TRUTH MEASUREMENT - hectares
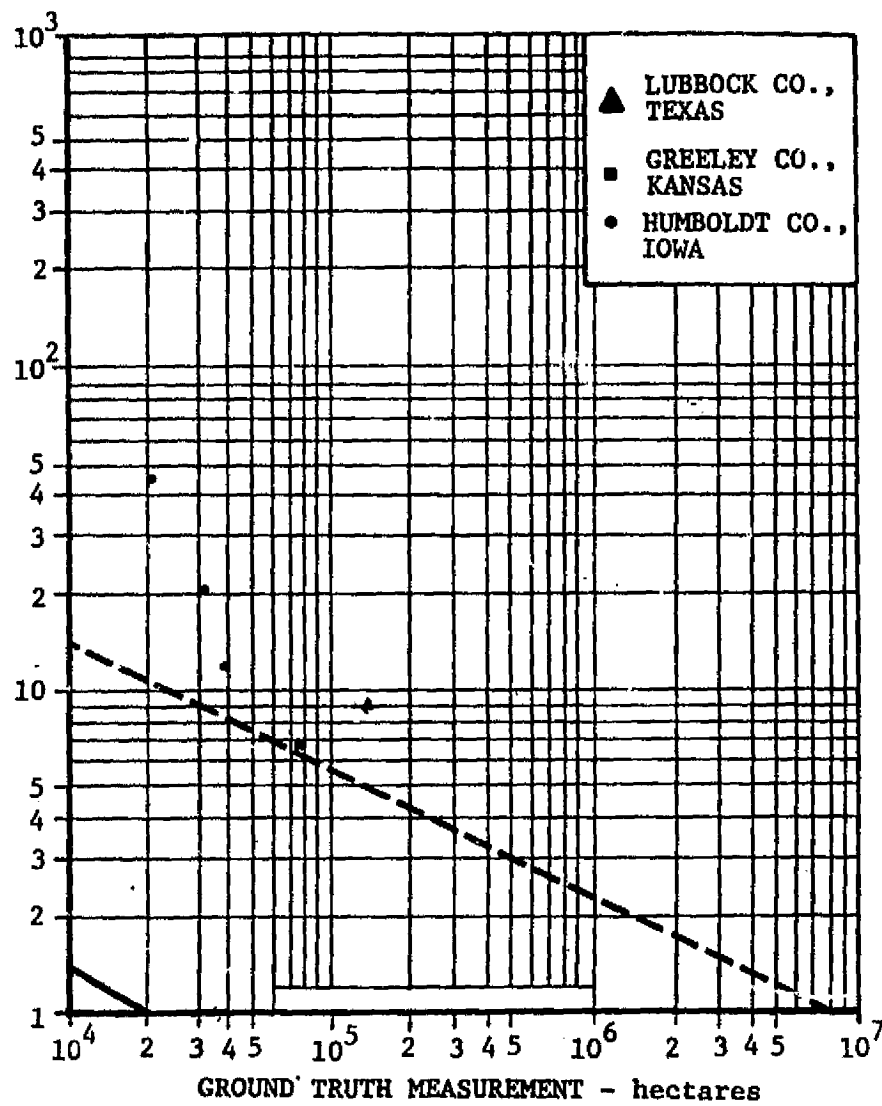
## TABLE C-10

### HIGHLIGHTS OF ERTS EXPERIMENT

Source: <u>A Study of the Utilization of ERTS-1 Data from the Wabash River Basin</u>, David A. Landgrebe and Staff, LARS, Purdue University, August 1974.

Scene Content:

| Illinois – Corn | 40% | | Indiana – Corn | 32% |
|---|---|---|---|---|
| Soybeans | 18% | | Soybeans | 25% |
| Grain Sorghum | | | Hay | |
| Alfalfa | | | Pasture | |
| Hay | | | Small Grain Stubble | 43% |
| Pasture | | | Woods | |
| Small Grain Stubble | 42% | | Urban | |
| Woods | | | | |
| Urban | | | | |

Time and Place of Measurement:

Illinois – 8/9/72
  DeKalb County
  Ogle County
  Lee County
(Analysis performed on total information from these three counties.)

Indiana – 8/21/73
| Lake County | Newton County |
|---|---|
| Porter County | White County |
| LaPorte County | Benton County |
| Starke County | Warren County |
| Pulaski County | Fountain County |
| Casper County | Tippecanoe County |

(Analysis performed on total information from these twelve counties.)  SEE ATTACHED MAP.

Method of Computation:

Illinois – Bias corrected and uncorrected ERTS results were determined, each on the basis of equal weights and non-equal weights, resulting in four acreage estimates made from ERTS data. The method used for unbiasing classification results involves multiplying the county classification results by the inverse of the test field classification performance matrix as follows:

$$A = CP^{-1}$$

Where:

    $A$ = 1 x n vector of crop acreages
    $C$ = classification vector with n crops or classes
    $P^{-1}$ = inverse of n x n classification performance matrix

Indiana – Acreage measurements based on pixel count classification.
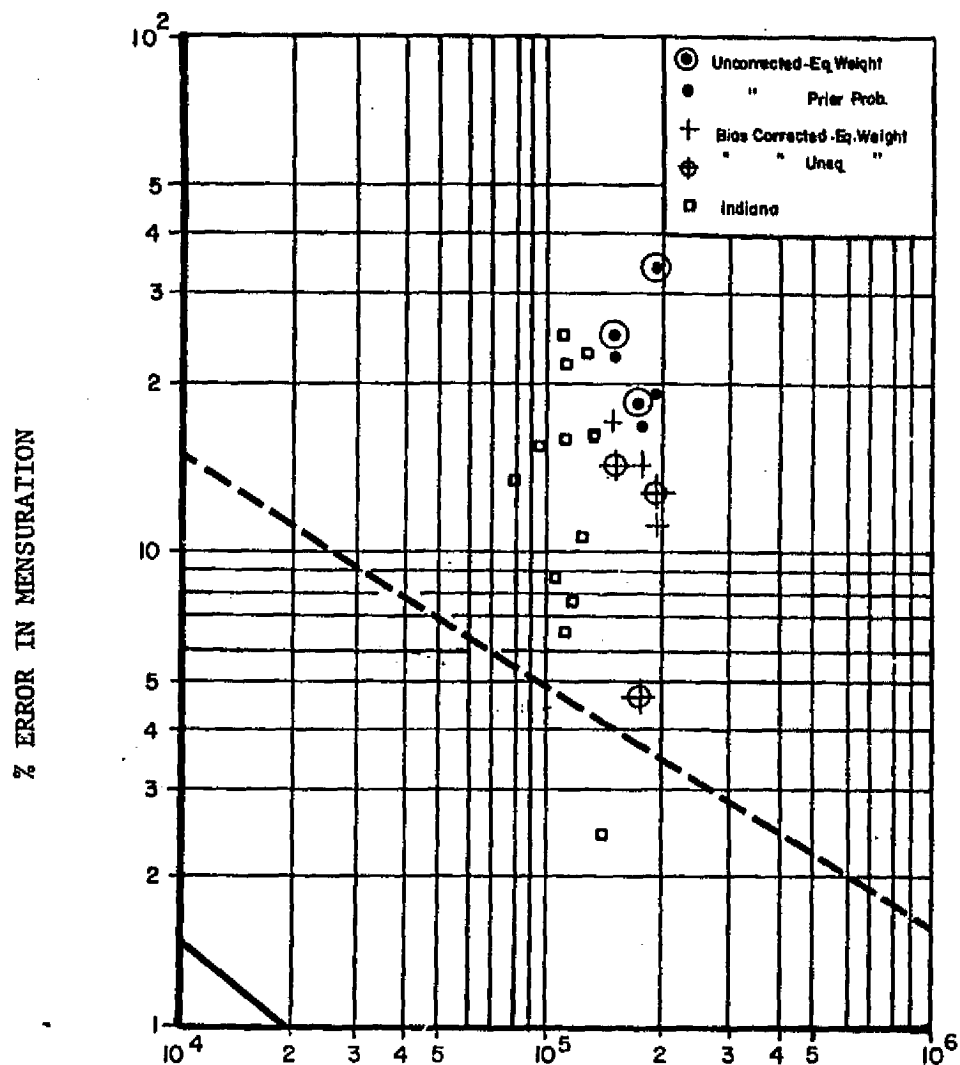
# FIGURE C-12

## EXPERIMENTAL ERRORS IN GEOMETRIC MENSURATION FROM LANDSAT IMAGERY – LANDGREBE

## TABLE C-11

## HIGHLIGHTS OF ERTS EXPERIMENT

Source: Identification and Quantification of Agriculture Crops, Deter-
mination of Dynamics of Land Use and Agricultural Census, Eng.
Humberto C. Garuti, Secretaria De Agricultura Y Ganaderia,
Argentina, February 1974.

Scene Content:

    Wheat - eared (flowering - grained)
    Barley
    Rye
    Corn - germinated
    Grain Sorghum - ready to plant
    Forage Sorghum - ready to plant
    Soya - planted
    Mijo - planted
    Sunflower - planted
    Alfalfa - shooted
    Mixed Pasture - shooted
    Natural Pasture - shooted
    Winter Cereal

Time and Place of Measurement:

Pampus Region of Argentina - 10/5/72

Method of Computation:

1. Manual Interpretation: Acreage estimates were obtained from 1:1,000,000
   amplified transparencies of false color images after classification.

2. Pixel Count: Ground truth available for only one scene; since the 3
   scenes used were successive frames from one ERTS pass, ground truth used
   as training data from first scene, then processed results from first
   scene used as initial training data for next scene, etc. (10% overlap
   between scenes.

# FIGURE C-13

## EXPERIMENTAL ERRORS IN GEOMETRIC MENSURATION FROM LANDSAT IMAGERY - GARUTI



GROUND TRUTH MEASUREMENT - hectares
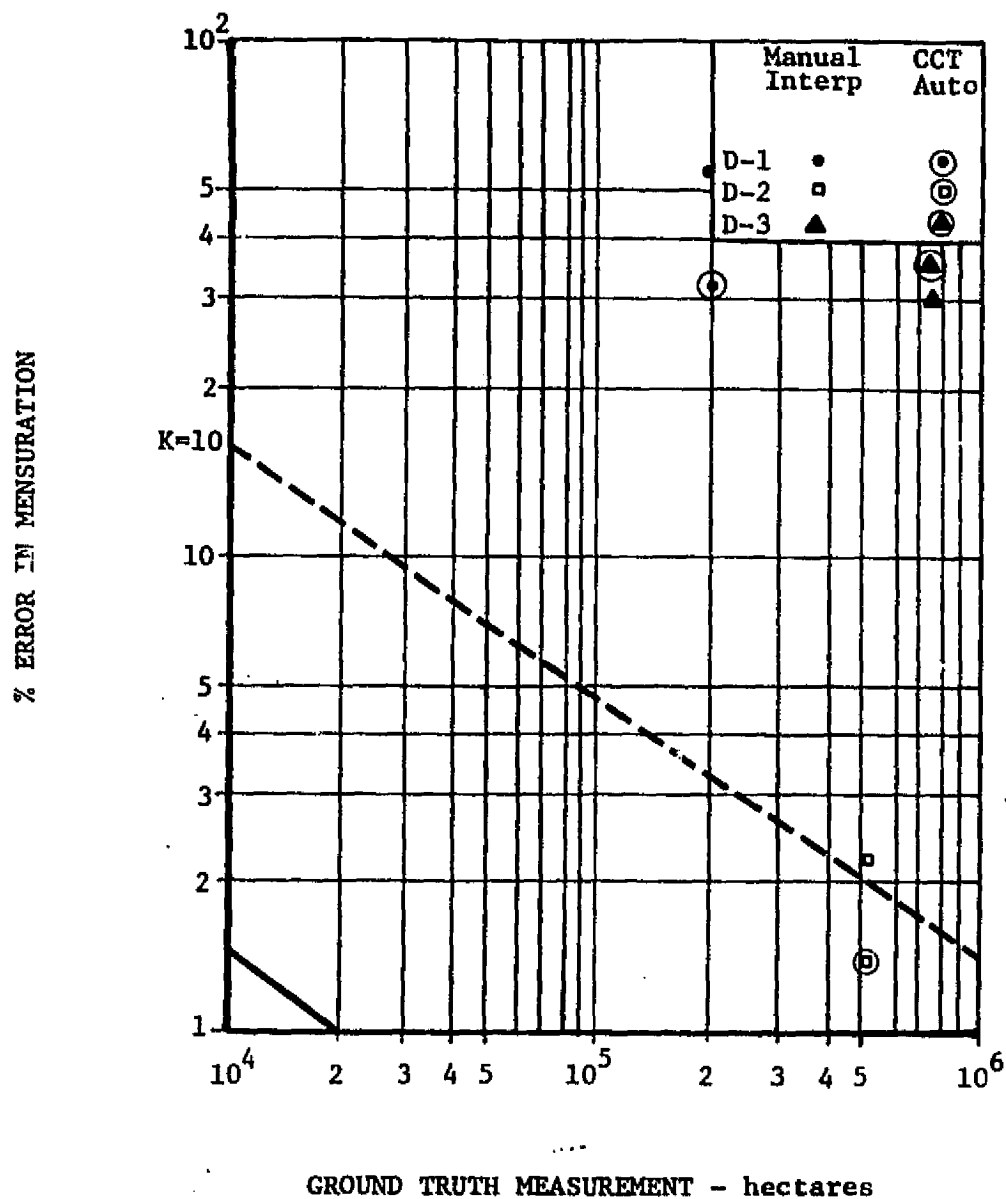
## Conclusions:

The best method for investigating large cultivated areas is automatic
discrimination, which assures speed and consideration of large areas.
A 96% safety factor was obtained in the general preliminary classifica-
tion.

## TABLE C-12

## HIGHLIGHTS OF ERTS EXPERIMENT

**Source:** Remote Sensing in Iowa Agriculture:  Identification and Classification of Iowa's Crops, Soils, and Forestry Resources Using ERTS-1 and Complimentary Underflight Imagery, J.P. Mahlstede et al., Iowa State University, May 1974.

**Scene Content:**

Corn
Alfalfa
Soybeans
Oats
Pasture
Water
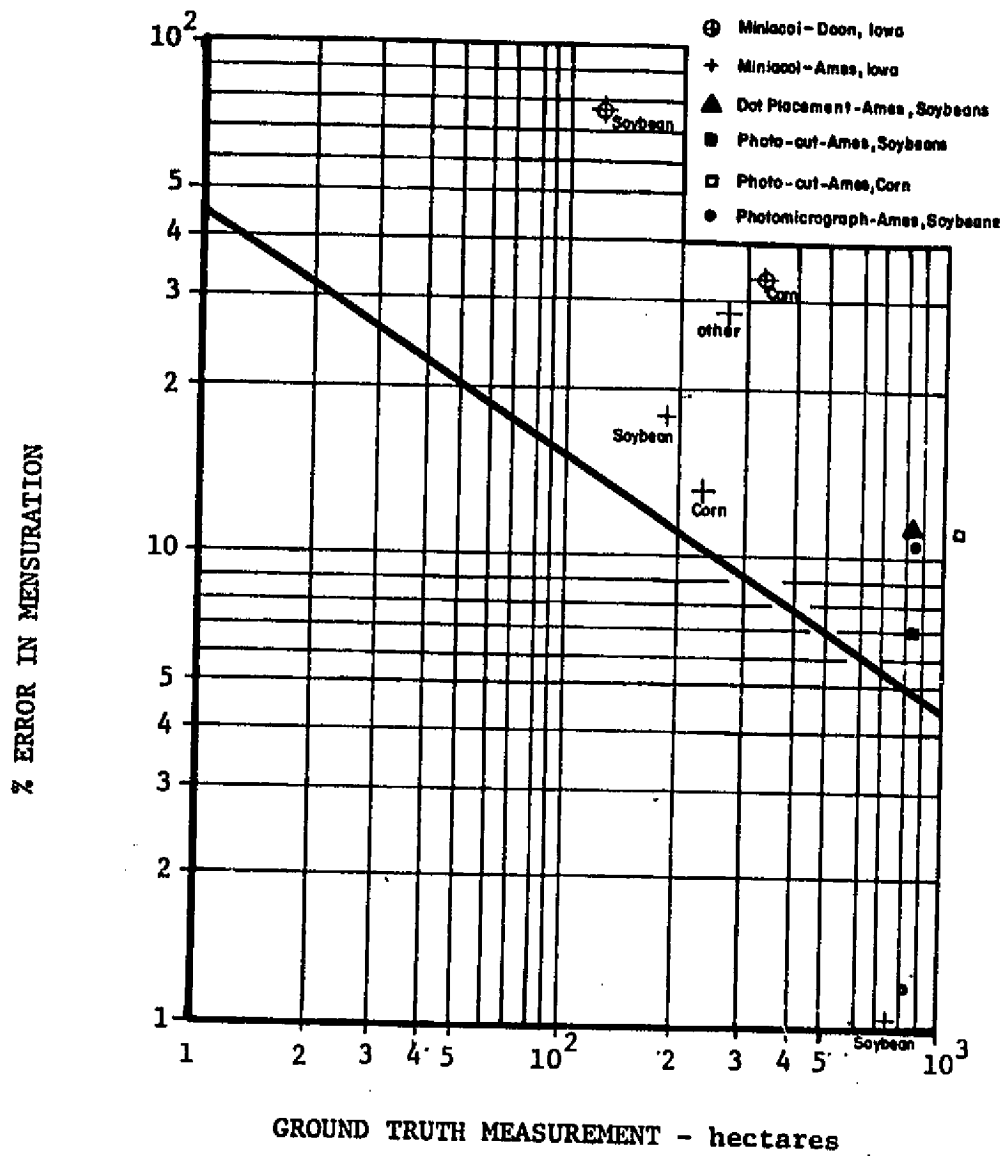Forest

**Time and Place of Measurement:**

Ames, Iowa - August 1972
Doon, Iowa - August 1972

**Method of Computation:**

Optical Processing

# FIGURE C-14

## EXPERIMENTAL ERRORS IN GEOMETRIC MENSURATION FROM LANDSAT IMAGERY - MAHLSTEDE

153

## TABLE C-13

## HIGHLIGHTS OF ERTS EXPERIMENT

Source:  EOS Thematic Mapper Technical Review, Ralph Shay et al., April 30 – May 2, 1975.

Scene Content:

Corn
Soybeans

Time and Place of Measurement:
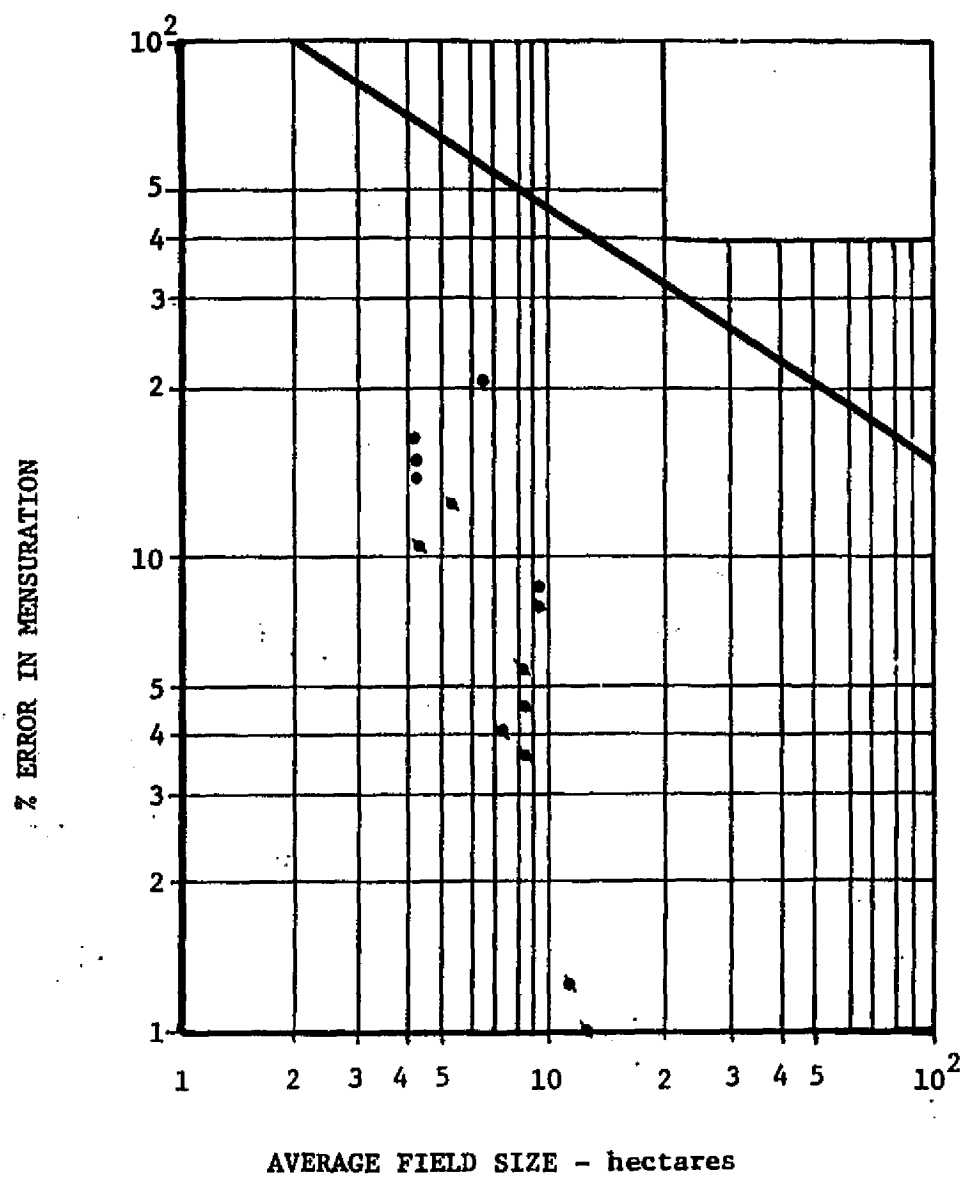
Midweat (Indiana and Illinois

Average Field Size:

6 - 17 Ha.

Method of Computation:

Pixel Count

FIGURE C-15

EXPERIMENTAL ERRORS IN GEOMETRIC MENSURATION FROM LANDSAT IMAGERY - SHAY(CITARS)



AVERAGE FIELD SIZE - hectares

## TABLE C-14

## HIGHLIGHTS OF ERTS EXPERIMENT

<u>Source</u>: <u>Regional Agricultural Surveys Using ERTS-1 Data</u>, Gene A. Thorley et al., Center for Remote Sensing Research, University of California, UN 640.

<u>Scene Content</u>:

| | |
|---|---|
| Range | Sugar Beets |
| Pasture | Walnut |
| Plowed Land | Fruit Trees |
| Grapes | Burned Areas |
| Alfalfa | Beans |
| Barley | Harvested |
| Sorghum | Cucumber |
| Tomato | |

<u>Time and Place of Measurement</u>:

San Joaquin County, California    July 26, 1972

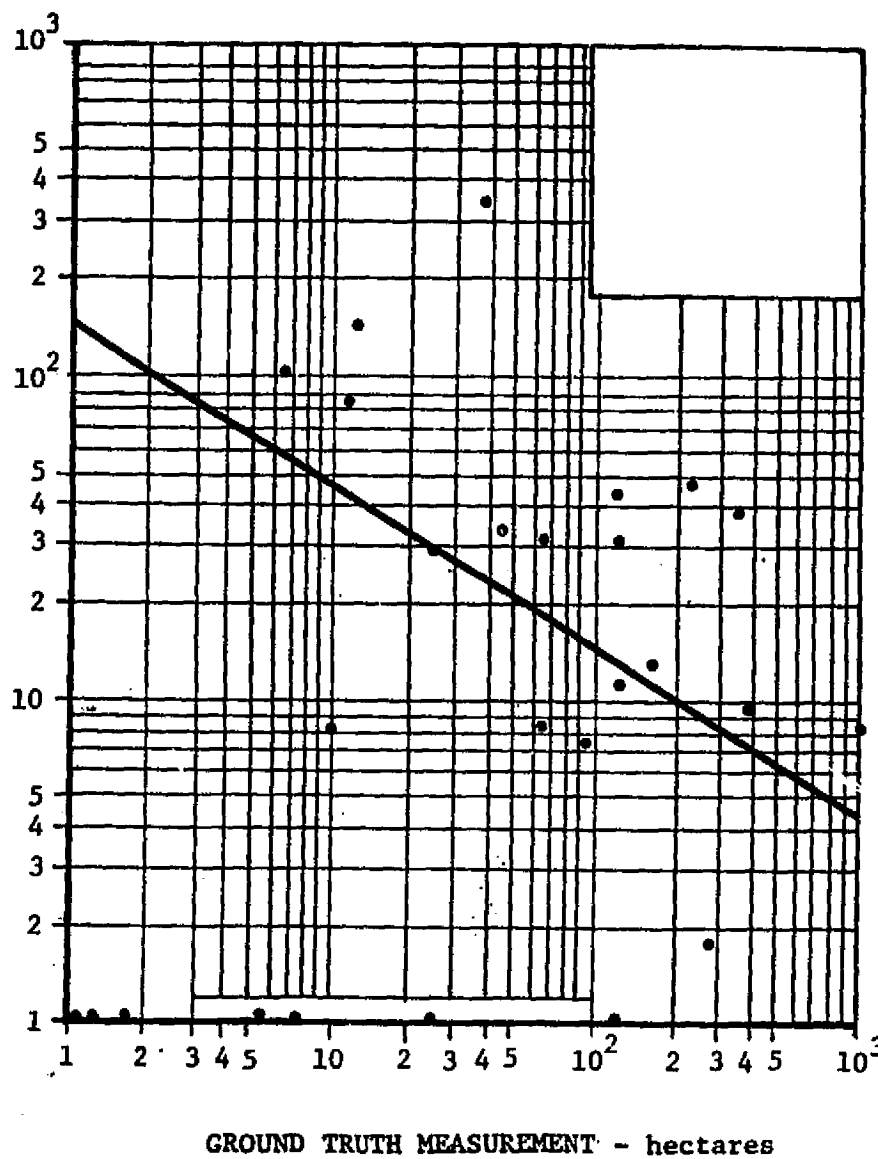<u>Average Field Size</u>:

Approximately 17 Ha.

<u>Method of Computation</u>:

Manual Interpretation
Automatic Interpretation

# FIGURE C-16

## EXPERIMENTAL ERRORS IN GEOMETRIC MENSURATION FROM LANDSAT IMAGERY – THORLEY



GROUND TRUTH MEASUREMENT – hectares

157

# APPENDIX D

## ANALYSIS OF AVAILABLE CROP SPECTRA

It has been known for a long time that laboratory spectra of single leaves exhibit a considerable degree of qualitative similarity. Personal communication by Professor R.N. Colwell of the University of California, Berkely, in 1975, who obtained numerous laboratory spectrograms, indicated that essentially "all leaves look alike." An example of this qualitative likeness is depicted in Figure D-1 which shows the qualitative similarities in the reflectances of single leafs in the spectral region from 0.5 to 2.5 μm of various plants: Bean, Avocado, Sorghum, Pigweed, Corn, Soybean, Orange, Peach, Cotton.

Leaf spectra however are not necessarily good indicators of the spectral reflectances of entire plants: especially since the available leaf spectra are taken "in vitro" at essentially normal incidences. One could speculate that specific patterns of leaves peculiar to each plant may emphasize preferred directions of reflection, or possibly even enhance one spectral region over others.

Analyses of spectra contained in the literature, see Table D-1, did qualitatively confirm the hypothesis of "similarity" of spectral shapes: however, the difficulties encountered in resolving discrepancies of calibration, stage of plant development, illumination and viewing angles, prohibited the formulation of general relationships valid between diverse plants, even between plants belonging to the same species but at differing stages of development, or subjected to different stresses.

158

# FIGURE D-1

### TYPICAL REFLECTANCE SPECTRA OF SINGLE LEAVES
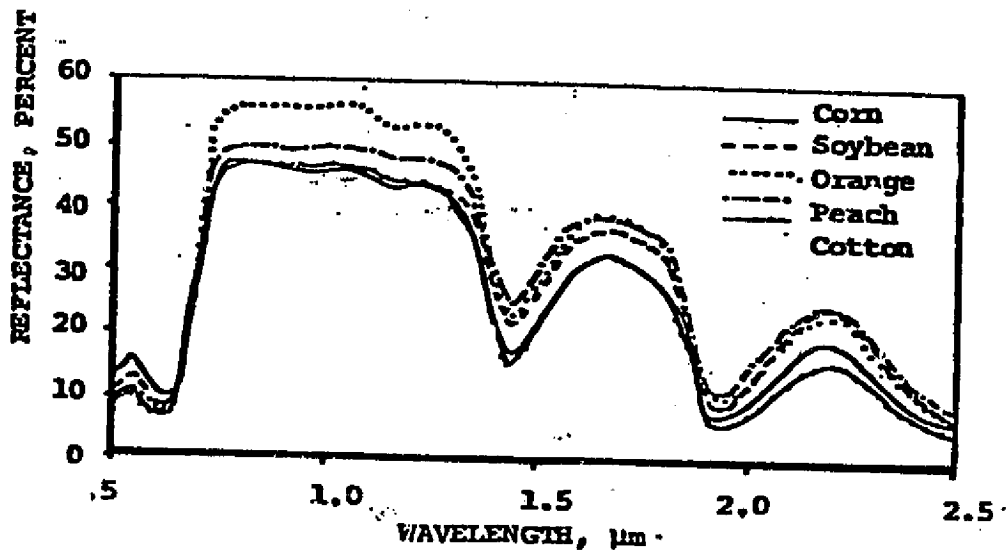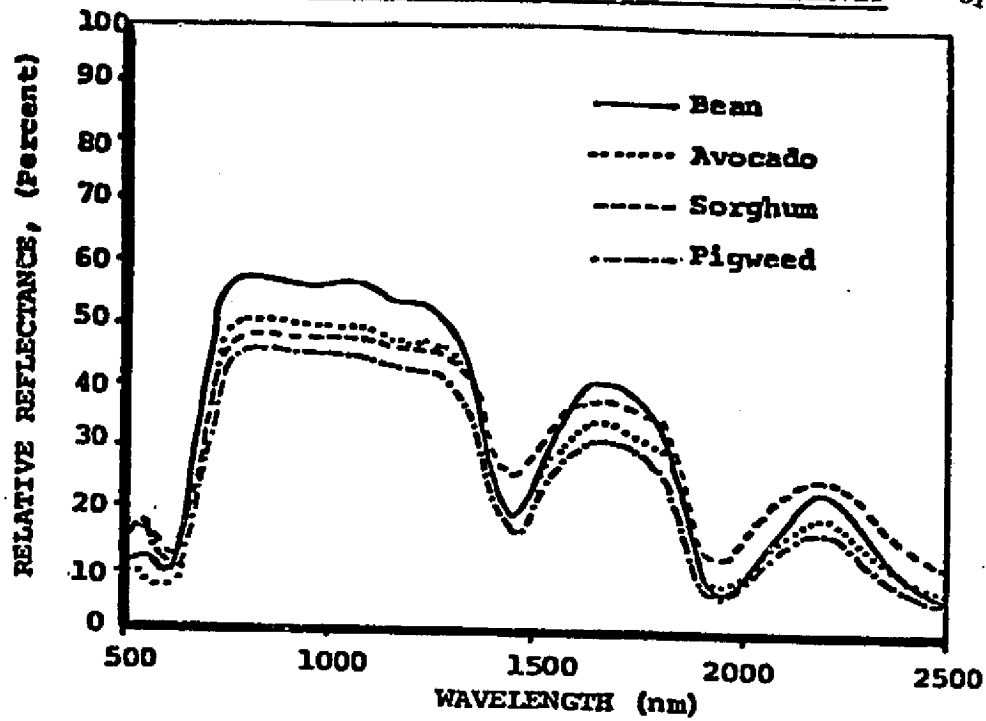


Gausman, Allen, Wiegand et.al., USDA-ARS Technical
Bulletin No. 1465, March 1973.

TABLE D-1

PRINCIPAL SPECTRAL DATA BASES AVAILABLE IN THE LITERATURE

| DESIGNATION | SPONSOR AGENCY | MEASUREMENT DATE | LOCATION |
|---|---|---|---|
| NTF | NASA/NTF | 12/73 - 5/74 | NTF - Bay St. Louis, MS |
| LARS | NASA | 7/68 | LARS - West Lafayette, IN |
| ERIM[*] | U.S. AIR FORCE | 1972 | NTIS - Springfield, VA |
| ERDL | U.S. ARMY | 1961 - 1962 | NTIS - Springfield, VA |
| WESLACO | USDA/ARS | 10/66 - 9/67 | USDA ARS - Weslaco, TX |
| KRINOV | USSR | 1936 - 1938 | NRC - Ottawa, Canada |
| IITRI[*] | NASA/USAF | Various | NTIS - Springfield, VA |

[*]Summary of data from other sources.

The resolution of this question had to await the advent of fine-grained, calibrated spectra.

Figures D-2 and D-3 present spectral radiance values produced by the Goddard Institute for Space Studies from a field program conducted in Imperial Valley, California. Figure D-2 shows spectra of the crops al-falfa, melons and sorghum at a 40 to 50% leaf cover. Figure D-3 shows the spectra of the crops wheat, alfalfa, sugar beets, sorghum, asparagus, each at an approximate leaf cover of 100%. These plots show that diverse crops at similar stages of development, i.e. at similar amounts of leaf cover, exhibit qualitatively similar spectral signatures.

SPECTRA OF VARIOUS CROPS AT 40-50% LEAF COVER IN IMPERIAL VALLEY, CALIFORNIA
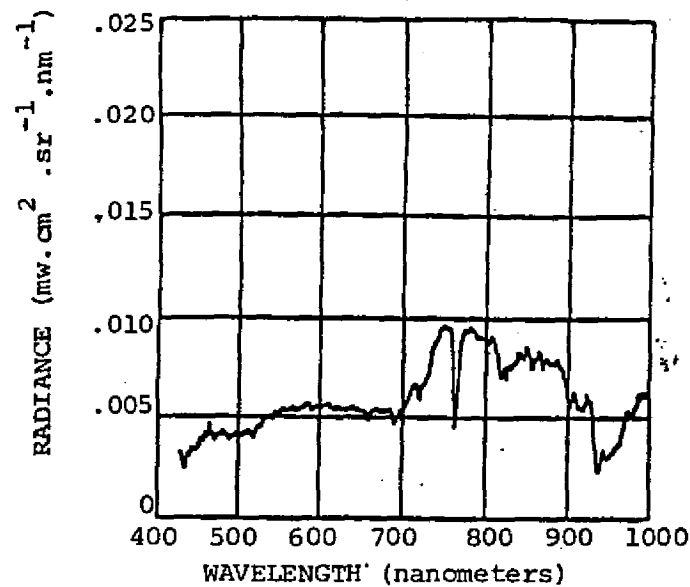


MELONS
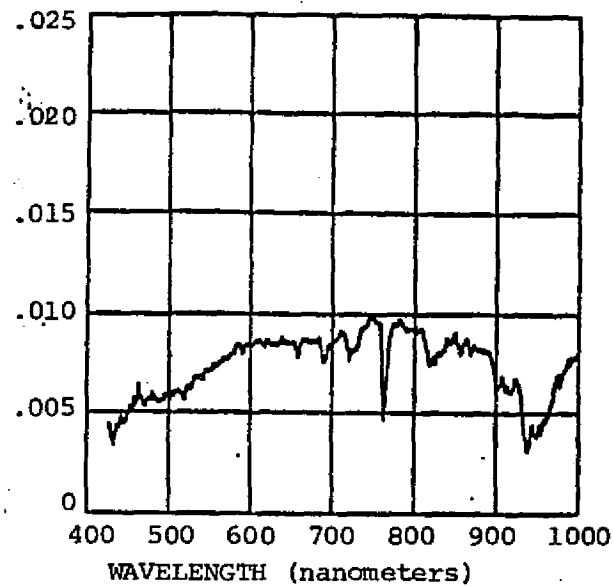
SORGHUM
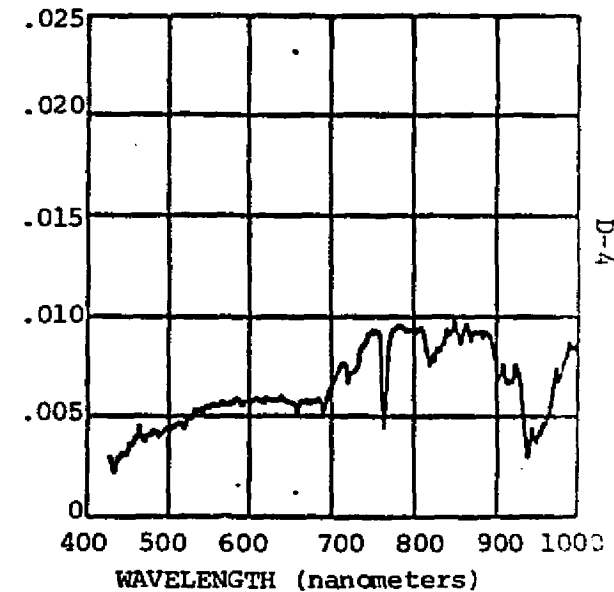
ALFALFA

SOURCE: Atlas of Selected Crop Spectra Imperial Valley, California
S.G. Ungar, NASA Institute for Space Studies, 1977

## SPECTRA OF VARIOUS CROPS AT 100% LEAF COVER IN IMPERIAL VALLEY, CALIFORNIA



WAVELENGTH (nanometers)
WHEAT

WAVELENGTH (nanometers)
SUGAR BEETS

WAVELENGTH (nanometers)
SORGHUM

WAVELENGTH (nanometers)
ASPARAGUS

WAVELENGTH (nanometers)
ALFALFA

RADIANCE $(\mu w \cdot cm^2 \cdot sr^{-1} \cdot nm^{-1})$

SOURCE: Atlas of Selected Crop Spectra Imperial Valley, California
S.G. Ungar, NASA Institute for Space Studies, 1977

D-5

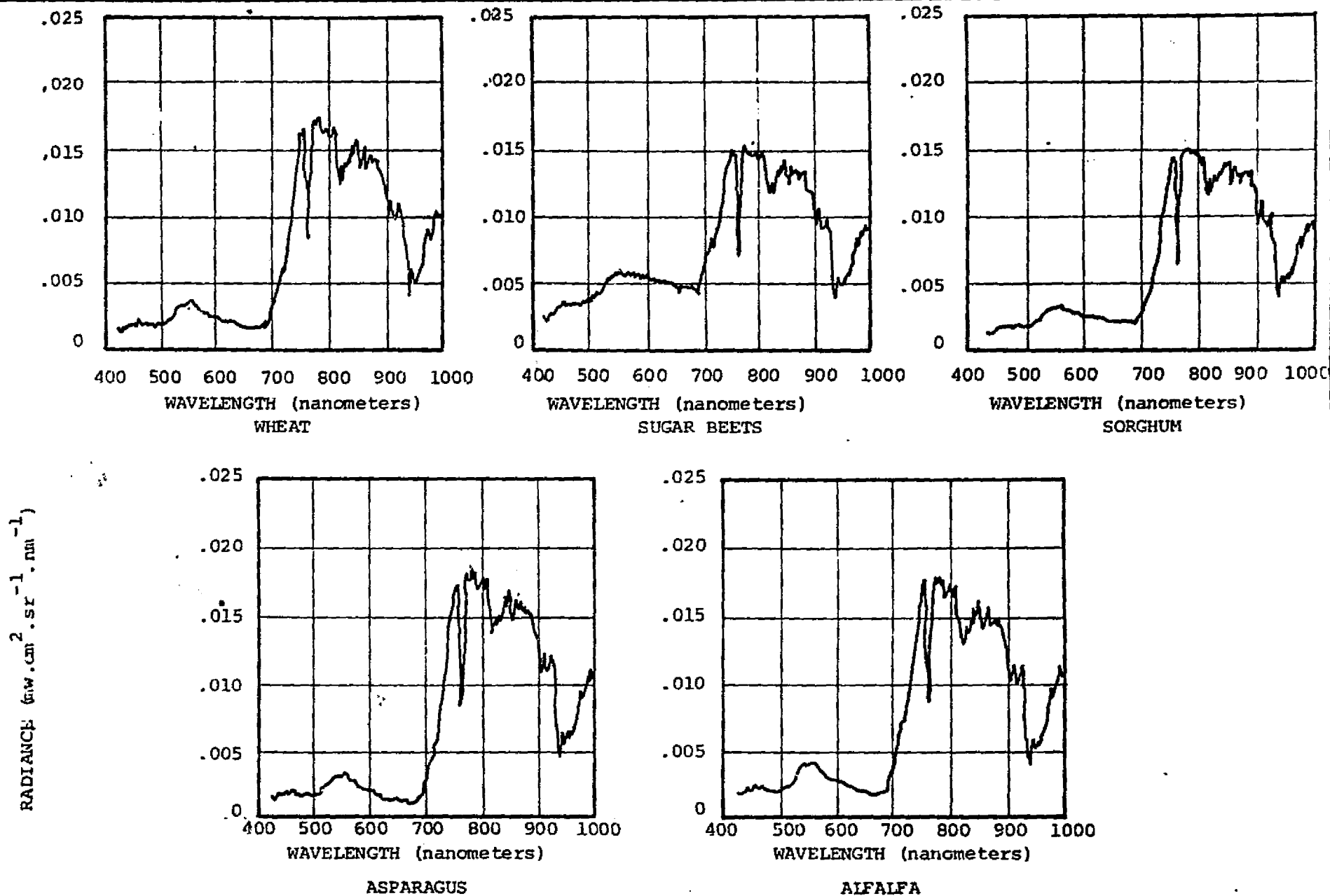A more quantitative determination of the spectral similarity between crops at like levels of leaf cover is obtained by examining the ratio between various spectral bands. For example, Table D-2 presents results of a correlation analysis between the spectral bands 0.7 to 0.8 µm and 0.8 to 0.9 µm for wheat and alfalfa crops with greater than 60% leaf cover.

### TABLE D-2

#### CORRELATION ANALYSIS BETWEEN BANDS 0.7-0.8 µm AND 0.8-0.9 µm

Number of Spectra Used: 149

| RESULTS: | COEFFICIENT OF CORRELATION | SLOPE | STD. ERROR OF SLOPE |
|---|---|---|---|
| WHEAT | 0.93 | 0.58 | 0.05 |
| ALFALFA | 0.97 | 0.58 | 0.02 |
| SOIL | 0.98 | 1.14 | 0.04 |

Source: ECOsystems analysis of a selected subset of spectra provided by Dr. S. Unger of the Goddard Institute for Space Science.

Each crop exhibits a "slope" (ratio of amplitude of band 0.8-0.9 to band 0.7-0.8 µm) of 0.58 and a high degree of correlation between the two spectral bands. The closeness of the slope values indicates that these crops generate similarly shaped spectra: whereas the high correlations indicate that the computed slope, hence the shape of the spectra is relatively stable for each crop at leaf cover areas greater than 60%. A number of such correlations were performed by ECOsystems with results similar to those obtained by other workers.

Table D-3 lists relative reflectances of irrigated wheat in the two spectral bands 0.74-0.76 µm and 0.78-0.80 µm and the ratio between the two

for the phenologic stages from early jointing through dead ripe for Bushland, Texas. [  ] These two bands show significant amplitude changes throughout the growth period. They exhibit 1 sigma variations of approximately 20% in the band 0.74-0.76 µm and 26% in the band 0.78-0.80 µm. The ratio between the two bands remains relatively constant, however, through this growth period, exhibiting a 1 sigma variation of approximately 8%. From these results it can be concluded that while the wheat field exhibits significant amplitude changes in mean reflectance throughout the later stages of the crop's growth, the shape of the spectra remains relatively constant.

TABLE D-3

RATIO OF RELATIVE REFLECTANCE 0.785 µm/0.745 µm FOR IRRIGATED

WHEAT FIELD NO. 9010W1, BUSHLAND, TX

| DATE | REFLECTANCE 0.74-0.76µm | REFLECTANCE 0.78-0.80µm | RATIO 1/2 | STAGE |
|------|------|------|------|------|
| 3/16 | .338 | .387 | 1.145 | Early Jointing |
|      | .331 | .382 | 1.154 | |
|      | .331 | .370 | 1.118 | |
|      | .315 | .337 | 1.070 | |
| 4/4 | .376 | .520 | 1.383 | Late Jointing |
|      | .256 | .299 | 1.168 | |
|      | .375 | .531 | 1.416 | |
|      | .393 | .546 | 1.389 | |
| 4/21 | .361 | .494 | 1.368 | Booting |
|      | .394 | .513 | 1.302 | |
|      | .387 | .498 | 1.287 | |
|      | .366 | .457 | 1.249 | |
| 5/19 | .304 | .382 | 1.257 | Early Ripe |
|      | .298 | .373 | 1.252 | |
|      | .316 | .379 | 1.199 | |
|      | .302 | .349 | 1.156 | |
| 5/27 | .248 | .304 | 1.226 | Late Ripe |
|      | .343 | .364 | 1.061 | |
|      | .229 | .245 | 1.070 | |
|      | .227 | .287 | 1.264 | |
| 6/4 | .231 | .269 | 1.165 | |
|      | .222 | .257 | 1.158 | |
|      | .227 | .261 | 1.150 | |
|      | .241 | .277 | 1.149 | |
| MEAN | .309 | .378 | 1.215 | |
| σ | .060 | .097 | .103 | |
| c.v. | .195 | .256 | .084 | |

This relationship appears to hold for other crops as well. Figures D-4 and D-5 present spectral plots of various crops throughout their growth for Hand County, South Dakota [ ] and Bushland, Texas. The similarity of shape, although with significant variations in amplitude from stage to stage are evident in these plots.

Figure D-6 plots single-date "snapshots" of spectra of various ground covers taken at the Bushland, Texas site. Note the emergence of irrigated wheat from the soil cover spectrum on the earlier date: its further growth, and the emergence of dryland wheat, on the later date. On the later date, whereas the spectral shapes of the emerging crops are qualitatively similar, their average magnitudes begin to differ significantly.

A striking example of the differences of amplitude displayed by crops at certain times during their growth cycle is shown in Figure D-7. Note the significant emergence of the crop rye over the others: its spectral shape, however, is qualitatively not very different from the spectral shapes of other crops. It should be noted that the April 4, 1974 curve for rye appeared upon investigation to be probably tainted by an imperfection in the experimental procedure---possibly an instrumental gain change of as much as 1.5:1. Even correcting for this possible mishap does however not modify the general relationships.

Figure D-8 plots the temporal progression of the mean reflectance of spring wheat, pasture, corn, oats, alfalfa and barley in the spectral band 0.76-0.9 μm for Hand County, South Dakota. It shows that temporal variations in mean reflectance differ from crop to crop.

## REFLECTANCE OF WINTER WHEAT AT DIFFERENT MATURITY STAGES

LOCATION: HAND COUNTY, SOUTH DAKOTA
SENSOR: FSS          DATE: 1975-76
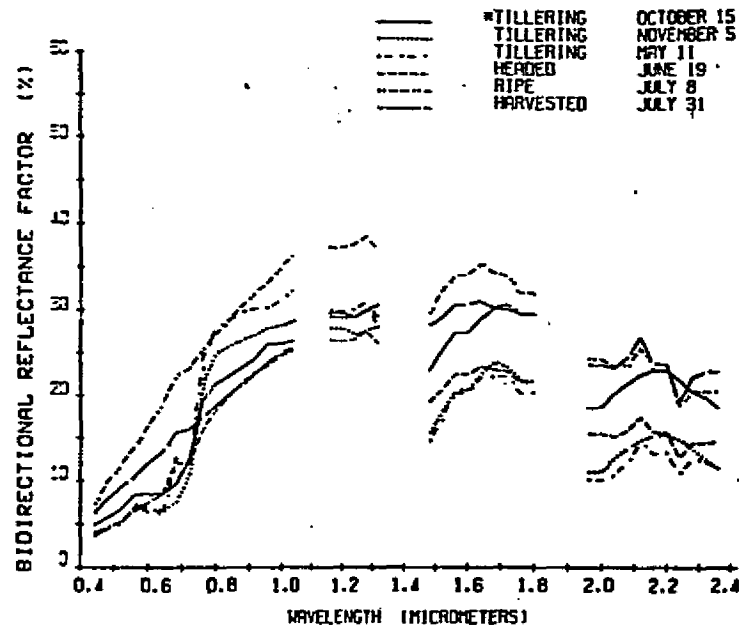
## REFLECTANCE OF SPRING WHEAT AT DIFFERENT MATURITY STAGES

LOCATION: HAND COUNTY, SOUTH DAKOTA
SENSOR: FSS          DATE: 1976



| *TILLERING | OCTOBER 15 |
| TILLERING | NOVEMBER 5 |
| TILLERING | MAY 11 |
| HEADED | JUNE 19 |
| RIPE | JULY 8 |
| HARVESTED | JULY 31 |



| * EMERGENCE | MAY 11 |
| HEADED | JUNE 19 |
| RIPE | JULY 8 |
| HARVESTED | JULY 31 |

175

166

PHENOLOGIC PROGRESSION OF THE SPECTRA OF IRRIGATED WHEAT

(BUSHLAND, TEXAS 1974)



PHENOLOGIC PROGRESSION OF THE SPECTRA OF IRRIGATED PASTURE

(BUSHLAND, TEXAS 1974)

## FIGURE D-6

### COMPARISON OF CROP SPECTRA FOR THE BUSHLAND, TEXAS SITE

APRIL 21, 1974

IRRIGATED WHEAT — DRYLAND WHEAT — SORGHUM — PASTURE

RELATIVE REFLECTANCE

WAVELENGTH



MAY 19, 1974

IRRIGATED WHEAT   DRYLAND WHEAT   PASTURE

RELATIVE REFLECTANCE

WAVELENGTH

ORIGINAL PAGE IS
OF POOR QUALITY

# COMPARISON OF CROP SPECTRA FOR THE BUSHLAND, TEXAS SITE



MARCH 16, 1974

RYE    IRRIGATED WHEAT    DRYLAND WHEAT    SORGHUM

RELATIVE REFLECTANCE

WAVELENGTH



APRIL 4, 1974

RYE    IRRIGATED WHEAT    DRYLAND WHEAT    SORGHUM

RELATIVE REFLECTANCE

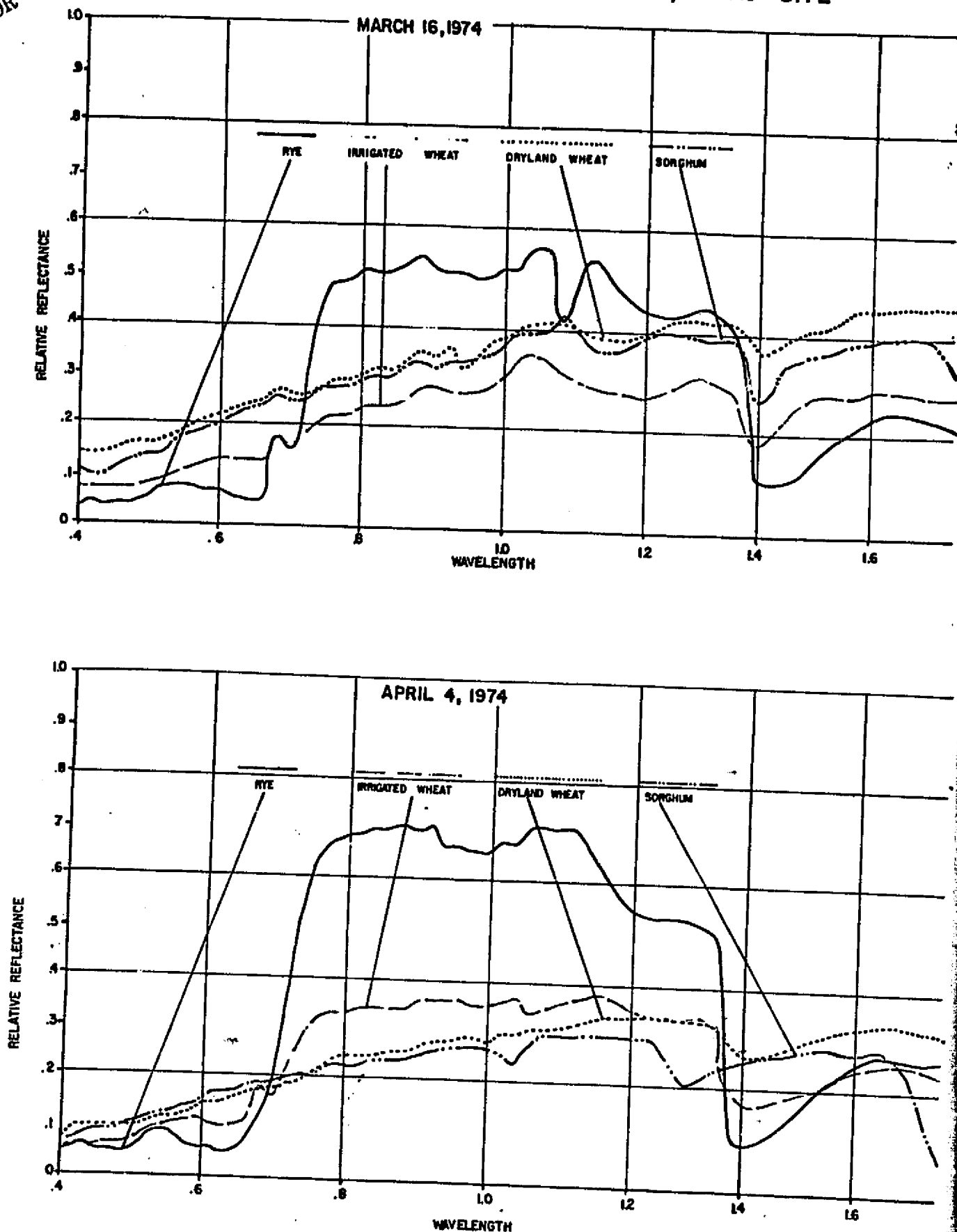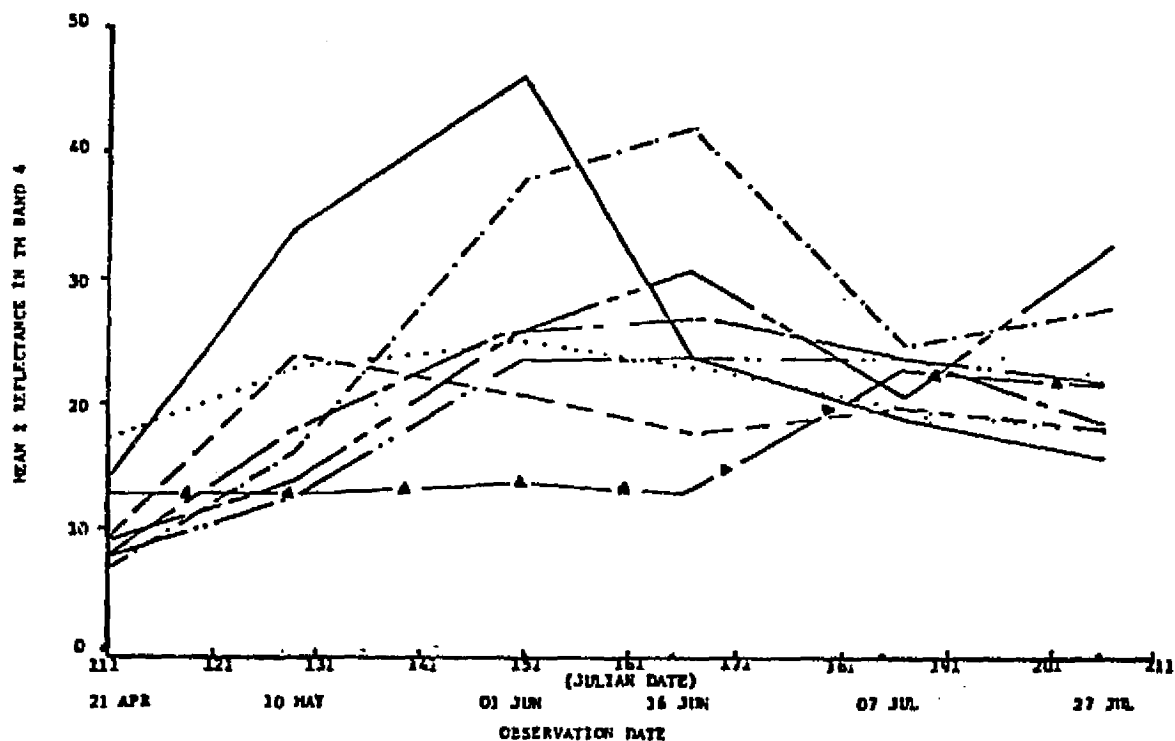WAVELENGTH

# FIGURE D-8

## TEMPORAL PROGRESSION OF THE MEAN REFLECTANCE FOR THE CROPS CONTAINED

## IN THE HAND COUNTY DATA SET

### UNADJUSTED DATA



| ——— — ——— | SPRING WHEAT 192 | ———▲——— | CORN 76 |
| ——————— — | SPRING WHEAT 169 | ——— · · ——— | OATS 136 |
| ——— — — — | WINTER WHEAT 168 | ——————— | ALFALFA |
| · · · · · · · | PASTURE 198 | —· —· —— | BARLEY |

It can be argued at this point that the findings thus far exemplified are peculiar to the limited number of regions within which they have been gathered: and that the hypothesis of "similarity" of spectral shapes is thereby not proven, but rather only inferrable as a working hypothesis.

There is of course no positive answer to this argument, whose thorough experimental verification would require an impracticable number of field data gathering programs. Heuristic arguments which point to the fact that the hypothesis of "similarity" is valid worldwide are:

1.  Personal communication by Dr. N.G. Kharin, Director, Remote Sensing Institute, Academy of Sciences of the Uzbekhistan Republic, USSR. Dr. Kharin confirmed in 1976 that from his findings in the USSR, spectra of trees and crops appear "very similar." His tabulations in [25] further confirm this fact.

2.  Spectra of rice gathered in hothouse environment by the French-Italian cooperative program "Agreste" show not only that various species of rice present "similar" reflectance spectra, but that these are essentially "similar" to the available spectra of other U.S. crops [26].

The mean reflectances of crops display fluctuations. Figure D-9 shows the fluctuations of mean reflectances measured among five spring wheat fields on July 28, 1976 in the Williams County, North Dakota LACIE Supersite. It also plots the results of 5 spectral scans of spring wheat field 334 located within this site: it shows that fluctuations occur not only among fields but within single fields. The Figures show that such fluctuations occur in all regions of the spectrum from 0.4 - 2.4 μm.

VARIABILITY IN REFLECTANCE
WITHIN A SPRING WHEAT FIELD

LOCATION: WILLIAMS COUNTY, NORTH DAKOTA
SENSOR: FSS          DATE: JULY 28, 1976
          FIELD: 334
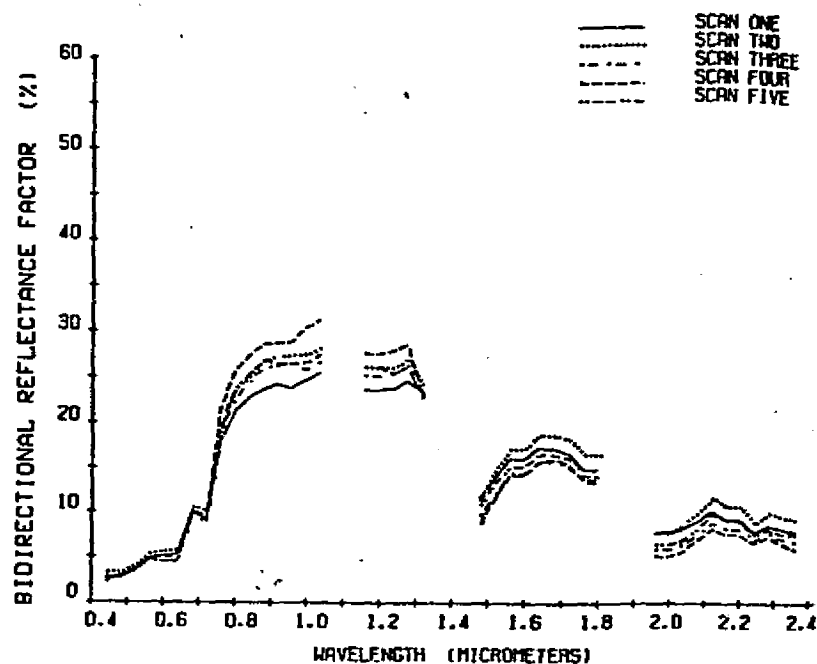
VARIABILITY IN REFLECTANCE
AMONG SPRING WHEAT FIELDS

LOCATION: WILLIAMS COUNTY, NORTH DAKOTA
SENSOR: FSS          DATE: JULY 28, 1976
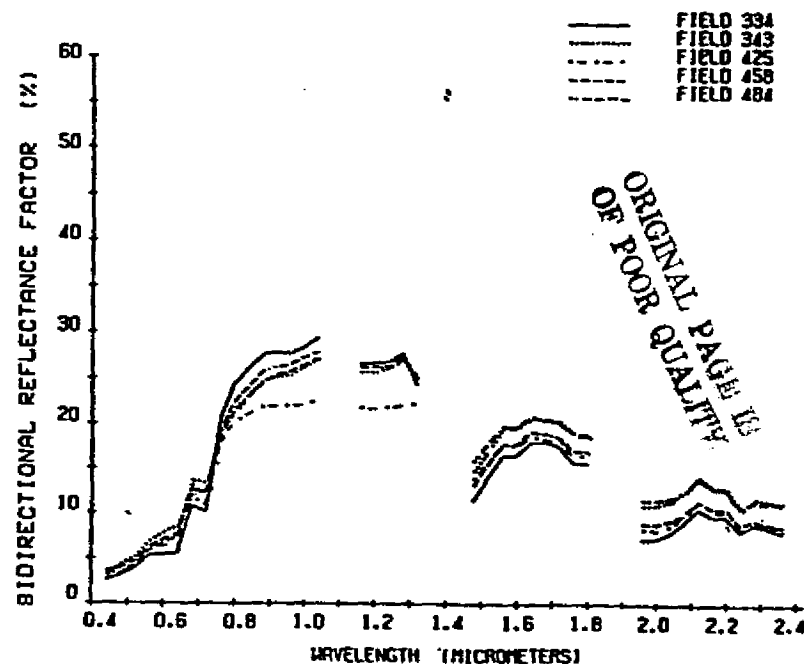
Figure D-10 plots coefficients of variation as a function of spectral wavelength for wheat in Bushland, Texas. It, and several other similar tests, indicates that an average spectral variability of order 10% is typical for reflectance fluctuations within a single field. Tables D-4 through D-6 list coefficients of variation for various multi-field crop classes for six dates in Hand County, South Dakota. They indicate that multi-field coefficients of variation are generally greater than 10%. The maximum multi-field reflectance fluctuation is exhibited by oats with a coefficient of variation of approximately 40% on June 16, 1977 in the spectral band from 2.08 to 2.35 μm. Typical coefficients of variation for multi-field crop classes appear to fall in the range from 10% to 30%.

FIGURE D-10

## COEFFICIENT OF VARIATION VERSUS WAVELENGTH

### BUSHLAND, TEXAS – MARCH 16, 1974
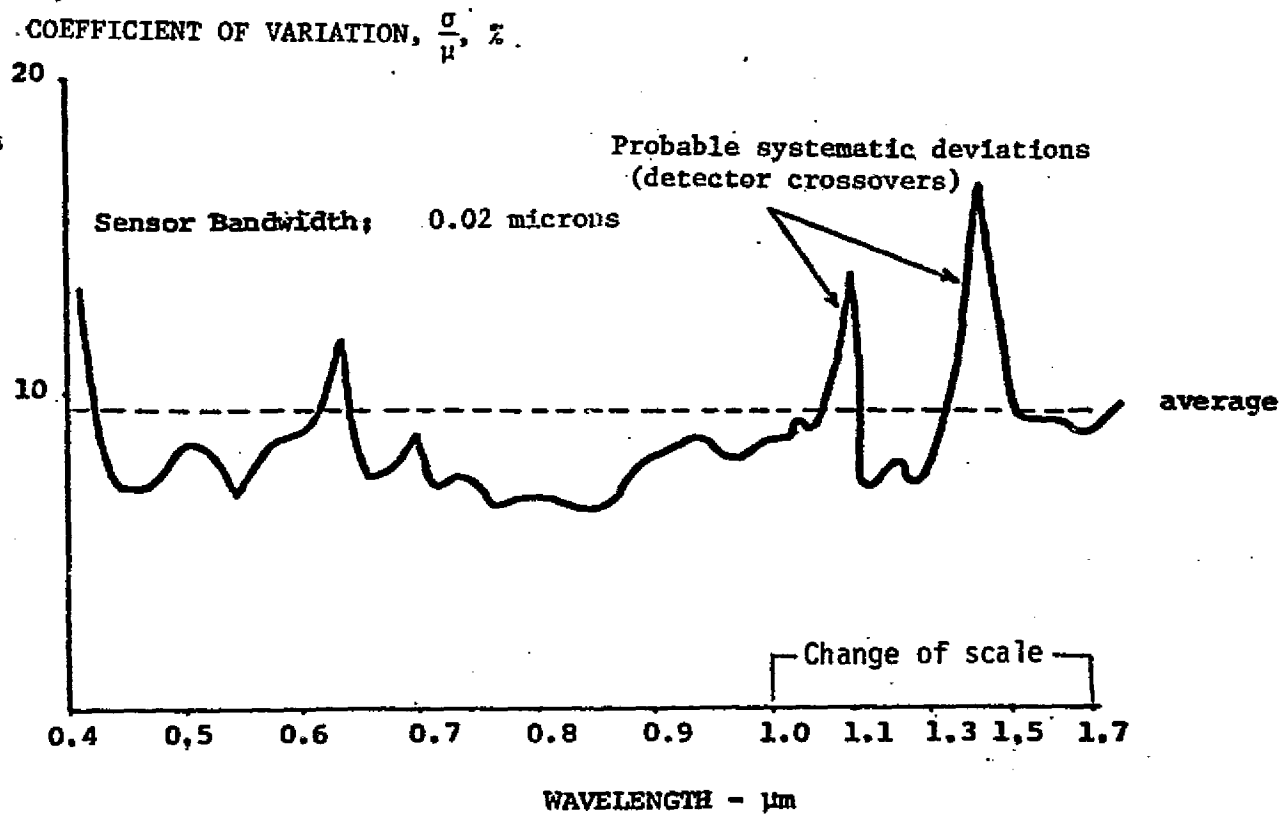
### SINGLE WHEAT FIELD

COEFFICIENT OF VARIATION, $\frac{\sigma}{\mu}$, %



WAVELENGTH – μm

## TABLE D-4

### COEFFICIENTS OF VARIATION OF REFLECTANCE VERSUS BAND AND DATE:  HAND COUNTY, S.D.

| SPECTRAL BAND | SPRING WHEAT CV% | | | | | | WINTER WHEAT CV% | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 4/21 | 5/10 | 6/01 | 6/16 | 7/07 | 7/27 | 4/21 | 5/10 | 6/01 | 6/16 | 7/07 | 7/27 |
| 0.450 - 0.520 | 13.5 | 7.1 | 13.6 | 15.8 | 15.0 | 15.0 | | | 20.7 | 33.9 | 19.3 | 22.0 |
| 0.520 - 0.600 | 15.5 | 6.6 | 9.4 | 14.0 | 14.2 | 16.9 | | | 14.6 | 31.1 | 17.2 | 18.4 |
| 0.630 - 0.680 | 13.9 | 6.4 | 9.9 | 14.7 | 16.2 | 18.0 | | | 14.4 | 34.0 | 15.6 | 16.2 |
| 0.760 - 0.900 | 13.6 | 6.7 | 8.5 | 13.6 | 12.0 | 16.6 | | | 13.5 | 29.0 | 7.6 | 14.2 |
| 1.550 - 1.750 | 6.1 | 3.5 | 9.2 | 16.0 | 10.7 | 7.1 | | | 19.7 | 39.7 | 15.0 | 12.6 |
| 2.080 - 2.350 | 7.3 | 4.8 | 17.5 | 25.8 | 19.4 | 8.5 | | | 34.1 | 48.4 | 20.5 | 21.4 |

AVERAGES OF 5 FIELDS, 74 TO 148 SPECTRA TAKEN AT EACH DATA

AVERAGES OF 4 FIELDS, 50 TO 109 SPECTRA TAKEN AT EACH DATE

## TABLE D-5

### COEFFICIENTS OF VARIATION OF REFLECTANCE VERSUS BAND AND DATE: HAND COUNTY, S.D.

| SPECTRAL BAND | CORN CV% | | | | | | OATS CV% | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 4/21 | 5/10 | 6/01 | 6/16 | 7/07 | 7/27 | 4/21 | 5/10 | 6/01 | 6/16 | 7/07 | 7/27 |
| 0.450 - 0.520 | 15.5 | 20.1 | 8.6 | 9.7 | 5.2 | 13.9 | 13.4 | 9.3 | 18.8 | 29.1 | 31.4 | 11.6 |
| 0.520 - 0.600 | 16.6 | 18.7 | 8.1 | 10.4 | 5.0 | 8.7 | 14.8 | 8.1 | 15.9 | 29.3 | 21.0 | 9.0 |
| 0.630 - 0.680 | 15.1 | 18.0 | 8.4 | 10.6 | 6.0 | 9.5 | 13.6 | 7.7 | 14.8 | 31.8 | 17.5 | 10.9 |
| 0.760 - 0.900 | 12.9 | 23.5 | 8.4 | 7.0 | 1.8 | 1.8 | 10.3 | 5.6 | 16.3 | 31.1 | 17.2 | 12.2 |
| 1.550 - 1.750 | 6.8 | 14.0 | 6.8 | 6.3 | 4.1 | 16.0 | 3.9 | 4.1 | 12.6 | 37.0 | 24.1 | 5.2 |
| 2.080 - 2.350 | 9.0 | 23.3 | 8.6 | 8.3 | 6.0 | 25.8 | 5.1 | 4.0 | 23.6 | 40.0 | 41.9 | 12.0 |

AVERAGE OF 3 FIELDS, 15 TO 64 SPECTRA TAKEN AT EACH DATE

AVERAGES OF 2 FIELDS, 57 TO 77 SPECTRA TAKEN AT EACH FIELD

# TABLE D-6

## COEFFICIENTS OF VARIATION OF REFLECTANCE VERSUS BAND DATE: HAND COUNTY, S.D.

| SPECTRAL BAND | ALFALFA CV% | | | | | | PASTURE CV% | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 4/21 | 5/10 | 6/01 | 6/16 | 7/07 | 7/27 | 4/21 | 5/10 | 6/01 | 6/16 | 7/07 | 7/27 |
| 0.450 - 0.520 | 8.5 | 24.6 | 7.4 | 28.4 | 5.3 | 16.4 | 15.3 | 15.2 | 17.7 | 20.8 | 15.0 | 13.4 |
| 0.520 - 0.600 | 10.8 | 16.3 | 15.7 | 27.0 | 4.8 | 13.4 | 12.3 | 11.7 | 11.5 | 15.6 | 11.3 | 10.1 |
| 0.630 - 0.680 | 9.1 | 14.6 | 5.3 | 29.3 | 5.3 | 12.4 | 11.1 | 11.0 | 10.6 | 15.1 | 10.5 | 9.7 |
| 0.760 - 0.900 | 20.6 | 17.5 | 6.4 | 48.2 | 2.7 | 4.1 | 11.5 | 9.2 | 12.7 | 21.8 | 9.7 | 9.3 |
| 1.550 - 1.750 | 4.5 | 15.9 | 2.3 | 38.7 | 3.7 | 12.9 | 13.3 | 10.2 | 14.1 | 18.7 | 10.1 | 10.1 |
| 2.080 - 2.350 | 8.3 | 36.8 | 3.7 | 50.7 | 5.1 | 19.5 | 15.0 | 16.1 | 24.1 | 26.8 | 15.7 | 14.9 |

AVERAGES OF 3 FIELDS, 26 TO 53 SPECTRA TAKEN AT EACH DATE

AVERAGES OF 4 FIELDS, 76 TO 142 SPECTRA TAKEN AT EACH DATE

176

## APPENDIX E

## INFORMATION EXTRACTION FROM RADIOMETRIC DATA

## DISCRIMINATION THEORY

There is a distinction between identification and discrimination. Discrimination means separation among species, i.e. their grouping into distinct classes, plus the assignment of a numerical value to each class: in our case, the total ground surface occupied by that class. The function of identification requires in addition the "naming" of each class, i.e. its assignment to predetermined categories.

The function of discrimination requires the sufficiently accurate recognition of distinctions between classes: identification requires in addition the identification of specific characteristics which univocally tag each distinct class as pertaining to a preassigned set.

The treatment which follows addresses the function of discrimination only, specifically: 1) the identification of the drivers of accuracy; and 2) the statement of their interrelationships.

### Single Band - Binary Mixture Case

Consider an area containing $N$ total pixels of two crop species: $N_A$ of species A and $N_B$ of species B. The reflectances of species A, B are characterized by Gaussian distributions with mean reflectances $\mu_A$, $\mu_B$ and standard deviations $\sigma_A$, $\sigma_B$.

The proportion of the area containing species A is:

$$\alpha = N_A/N$$

That containing species B is:

$$(1 - \alpha) = N_B/N$$

The probability density function p (x) of the distribution of the binary crop mixture is:

$$p\ (x) = \alpha\ P_A\ (x) + (1-\alpha)\ P_B(x) \qquad (E-1)$$

where $P_A$ is the conditional probability of x belonging to Class A and $P_B$ is the conditional probability of x belonging to Class B, and $\alpha$ is the a priori probability of n belonging to Class A.

The expected mean reflectance $\mu_\alpha$ of the N pixels is given by:

$$\mu_\alpha = \int_{-\infty}^{\infty} x\ p\ (x)\ dx =$$

$$= \int_{-\infty}^{\infty} x\ [\alpha\ P_A\ (x) + (1-\alpha)\ P_B\ (x)]\ dx =$$

$$= \alpha\ \mu_A + (1-\alpha)\ \mu_B \qquad (E-2)$$

and the expected variance by:

$$\sigma_\alpha^2 = \int_{-\infty}^{\infty} x^2 \, p(x) \, dx =$$

$$= \int_{-\infty}^{\infty} x^2 \, [\alpha \, P_A \, (x) + (1-\alpha) \, P_B \, (x)] \, dx =$$

$$= \alpha \sigma_A^2 + (1-\alpha) \, \sigma_B^2 + \alpha \, (1-\alpha) \, (\mu_B - \mu_A)^2 \qquad \text{(E-3)}$$

For sufficiently large samples the sample mean M tends to a normal variable with mean $\mu_\alpha$ and standard deviation $\sigma_\alpha / \sqrt{N}$. Using the standard definition of the q% probability that the true value is contained in an interval

$$M - k \, (\frac{\sigma_\alpha}{\sqrt{N}}) < \mu_\alpha < M + k \, (\frac{\sigma_\alpha}{\sqrt{N}})$$

where t is a constant whose values are given below:

| CONFIDENCE LEVEL q% | LEVEL OF SIGNIFICANCE (100-q)% | k |
|---|---|---|
| 90 | 10 | 1.64 |
| 95 | 5 | 1.96 |
| 99 | 1 | 2.58 |
| 99.9 | 0.1 | 3.29 |
| 99.99 | 0.01 | 3.89 |

For species A with $\hat{\mu}_A$ as the estimate of the mean value for species A, the true value $\mu_A$ lies between the limits

$$\tilde{\mu}_A - \frac{k \, \sigma_A}{\sqrt{N_A}} \; < \mu_A << \tilde{\mu}_A + \frac{k \, \sigma_A}{\sqrt{N_A}} \qquad\qquad (E\text{-}4)$$

at the confidence level q. Similarly for species B:

$$\tilde{\mu}_B - \frac{k \, \sigma_B}{\sqrt{N_B}} \; < \mu_B < \tilde{\mu}_B + \frac{k \, \sigma_B}{\sqrt{N_B}} \qquad\qquad (E\text{-}5)$$

Adding the two equations and weighting them by their a priori proportions, yields:

$$\alpha\tilde{\mu}_A + (1\text{-}\alpha)\, \tilde{\mu}_B - k \left[\alpha \frac{\sigma_A}{\sqrt{N_A}} + (1\text{-}\alpha) \frac{\sigma_B}{\sqrt{N_B}}\right] \; < \alpha\mu_A + (1\text{-}\alpha)\, \mu_B <$$

$$< \alpha\tilde{\mu}_A + (1\text{-}\alpha)\, \tilde{\mu}_B + k \left[\alpha \frac{\sigma_A}{\sqrt{N_A}} + (1\text{-}\alpha) \frac{\sigma_B}{\sqrt{N_B}}\right] \qquad\qquad (E\text{-}6)$$

Rewriting the equation:

$$M - k \left[\alpha \frac{\sigma_A}{\sqrt{N_A}} + \frac{(1\text{-}\alpha)\, \sigma_B}{\sqrt{N_B}}\right] < \alpha\, (\mu_A\text{-}\mu_B) + \mu_B <$$

$$< M + k \left[\alpha \frac{\sigma_A}{\sqrt{N_A}} + \frac{(1\text{-}\alpha)\, \sigma_B}{\sqrt{N_B}}\right] \qquad\qquad (E\text{-}7)$$

Where:

$$M = \alpha \, \tilde{\mu}_A + (1-\alpha) \, \tilde{\mu}_B$$

Introducing the definitions below in (E-7):

$$a = \frac{\mu_B - \mu_A}{\sigma_A}$$

$$b = \frac{\sigma_B}{\sigma_A}$$

$$\tilde{\alpha} = \frac{\mu_B - M}{a \, \sigma_A}$$

yields:

$$\tilde{\alpha} - \frac{k}{a \sqrt{N}} \; [\sqrt{\alpha} + b \sqrt{1-\alpha}\,] < \alpha$$

$$\lesssim \tilde{\alpha} + \frac{k}{a\sqrt{N}} \; [\sqrt{\alpha} + b \sqrt{1-\alpha} \qquad \text{(E-8)}$$

Thus the absolute value of the error e of the estimate of $\alpha$ is bounded by:

$$e < \frac{k}{a \sqrt{N}} \; [\sqrt{\alpha} + b \; \sqrt{1-\alpha}\,] =$$

$$= \frac{1}{\alpha \sqrt{N}} \; \frac{k}{\mu_A - \mu_B} \sqrt{\alpha \, \sigma_A^2 + (1-\alpha) \, \sigma_b^2} \qquad \text{(E-9)}$$

# REFERENCES

1. The Pratical Utilization of Remote Sensing Technology for the Management and Conscrvation of Natural Resources, Part 1: Crop Forecasting, P. Castruccio & H. Loats, Jr., United Nations. PO #5-20-255, March 1974.

2. Assessment of the World Food Situation – Present and Future, United Nations World Food Conference, Rome 5-16 November 1974.

3. Agricultural Statistics 1976 & Agricultural Statistics 1977, U.S. Department of Agriculture.

4. The State of Food and Agriculture 1976, Food and Agriculture Organization of the United Nations, Rome, 1975.

5. Letter by Willam Kibler, Administration, STS, dated February 19, 1976.

6. Personal Communications from Mr. E. Mercati, Director, Agricultural Section, IstiutoCentrale di Statistica, Rome, Italy.

7. Canada Year Book, 1975, 1971 Census of Canada – Agriculture Canada, Agriculture – Nova Scotia, Agriculture – New Brunswick, Agriculture Prince Edward Island, & Agriculture – Newfoundland.

8. Global Crop Production Forecasting An, Analysis of the Data System Problems and Solutions, Final Report, Contract NAS8-32491, May 1978, Report Number 78HV301 General Electric Company.

9. Global Agricultural Statistics in Support of Sigma Squared. P. Castruccio H. Loats, T. Fowler, & A. Baumann, January 6, 1977 ECOsystems Int'l. Inc.

10. U.S. Weekly Crop and Weather Bulletin 1977.

11. Time Study Final Report: Volume I – Results of Temporal Investigations 20 October 1978 (Draft) Contract No. NAS5-23412, Mod 149 General Electric Company and ECOsystems International.

12. Agricultural Commodities – Projections for 1975 and 1985, Volume I, Part I, Part II, Part III.

    Agricultural Commodities – Projections for 1975-1986, Volume II Methodological Notes, Statistical Appendix.

13. National Methods of Collecting Agricultural Statistics, Volume I: Asia and the Far East, European and North American Regions

    Volume II: African, Latin American and Near East Regions.

14. Personal Communications from Mr. Dale Rowe, GE Missile & Space Division, Huntsville Operations, Huntsville, Alabama

15. Ungar GISS Report Spectra

16. The ERTS-1 Investigation (ER-600) Volume I - ERTS Agricultural Analysis, R. Bryan Erb. NASA, August 1974.

17. Reflectance of Vegetation, Soil and Water, Craig L. Wiegand et al., Agricultural Research Service, USDA, November 1974.

18. Evaluation and Comparison of ERTS Measurements of Major Crops and Soil Associations for Selected Test Sites in the Central United States, Marion F. Baumgardner and Staff, LARS, Purdue University, March 1974.

19. A Study of the Utilization of ERTS-1 Data from the Wabash River Basin, David A. Landgrebe and Staff, LARS, Purdue University August 1974.

20. Identification and Quantification of Agriculture Crops, Determination of Dynamics of Land Use and Agricultural Census, Eng. Humberto, C. Garuti, Secretaria De Agricultura Y Ganaderia, Argentina, February 1974.

21. Remote Sensing in Iowa Agriculture: Identification and Classification of Iowa's Crops, Soils, and Forestry Resources Using ERTS-1 and Complimentary Underflight Imagery, J.P. Mahlstede et al., Iowa State Univeristy, May 1974.

22. EOS Thematic Mapper Technical Review, Ralph Shay et al., April 30- May 2, 1975.

23. Regional Agricultural Surveys Using ERTS-Data, Gene A. Thorley et al. Center for Remote Sensing Research, University of California, UN 640.

24. Distanzionnie Metodi Jzuchenia Rastitelnosti, N.G. Kharin, Academy of Sciences of the USSR and the Turkmenistan Republic, Moscow, 1975.

25. Agreste Program Part 2: Italian Test Sites Progress Report, Mar. - Jun 1976 (Commission of the European Communities)